DOCUMENT RESUME

ED 043 122 | 24 | EA 003 071

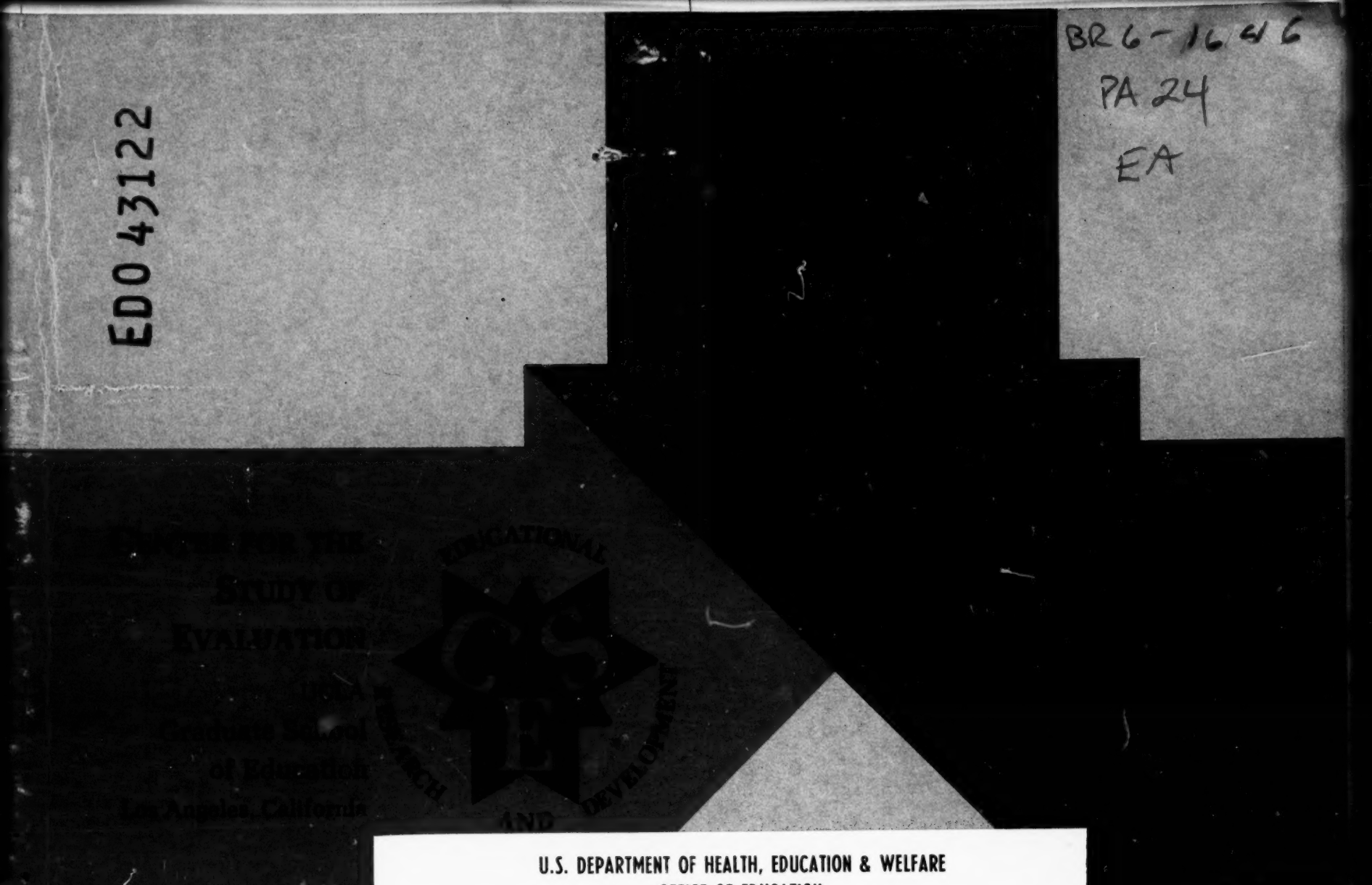
AUTHOR	Alkin, Marvin C.
TITLE	Products for Improving Educational Evaluation. Center for the Study of Evaluation. Fifth Annual Report.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Department of Health, Education, and Welfare, Washington, D.C. National Center for Educational Research and Development.
BUREAU NO	BR-6-1646
PUB DATE	1 Aug 70
CONTRACT	OEC-4-6-061646-1909
NOTE	56p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.90
DESCRIPTORS	*Evaluation Methods, *Evaluation Techniques, *Research and Development Centers

ABSTRACT

This publication provides background information on the functions and operations of the Center for the Study of Evaluation and reports on such center products as Insructional Objectives Exchange (IOX), CSE Elementary School Test Evaluations, and Evaluation Workshop I. Appendixes include: a summary of center accomplishments; a list of the center's research staff; a publication list; a breakdown of operating funds for the fiscal year 1970, percentage allocation by project; and a chart of the administrative structure. (LLR)

ED043122

BR6-1646
PA 24
EA



U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PRODUCTS FOR IMPROVING EDUCATIONAL EVALUATION

Fifth Annual Report
to the U. S. Office of Education

August 1, 1970

EA 003 071

+8

UNIVERSITY OF CALIFORNIA, LOS ANGELES

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

CENTER FOR THE STUDY OF EVALUATION
UCLA GRADUATE SCHOOL OF EDUCATION
LOS ANGELES, CALIFORNIA 90024

ED043122

August 1, 1970

Division of Educational Laboratories
National Center for Educational
Research and Development
United States Office of Education
Washington, D. C. 20202

Gentlemen:

Enclosed is the Fifth Annual Report of the Center for the Study of Evaluation. In this report we have attempted not only to summarize our activities and accomplishments of the past year, but to do so in a lucid, easily readable manner that would be appropriate for broad scale dissemination to audiences of various types.

As you will note from the materials that follow, the Center has enjoyed a great number of significant accomplishments during the past year and has now produced a number of products of considerable importance to the improvement of evaluation practice.

We look forward to continued service to American education.

Sincerely yours,

Marvin C Alkin

Marvin C. Alkin
Director

PRODUCTS FOR IMPROVING EDUCATIONAL EVALUATION

Fifth Annual Report
of the
Center for the Study
of Evaluation

Marvin C. Alkin, Director

August 1, 1970

UCLA
Graduate School of Education
Los Angeles, California

U. S. Office of Education
Government Contract:
OEC4-6-061646-1909

UCLA CSE Evaluation comment

September, 1970 Vol. 2, No. 3

Center for the Study of Evaluation

PRODUCTS FOR IMPROVING EDUCATIONAL EVALUATION

Fifth Annual Report to the U.S. Office of Education

Marvin C. Alkin, Director

Each year a total of 4.5 billion dollars is spent by the Federal Government on education, and an estimated 40.6 billion dollars by all public schools in the United States. There is an urgent need to develop procedures which will help to assure that dollars for education are being spent wisely. There is an urgent need to develop systems and procedures for improving the evaluation of education. There is an urgent need to develop procedures for evaluating the effectiveness of various instructional programs.

The mission of the Center for the Study of Evaluation (CSE), established in 1966, is to produce new materials, practices, and knowledge leading to the development of systems for evaluating education which can be adopted and implemented by educational agencies. The scope of activities at the Center includes the following: (1) the development of procedures and methodologies needed in the practical conduct of evaluation studies of various types; and (2) the development of generalizable concepts of evaluation relevant to different levels of education.

In addition to the wide variety of CSE technical publications which have had a far reaching effect upon the work of other educational researchers and developers, the Center has developed a number of important products. Among the most notable of these are:

The Instructional Objectives Exchange (IOX), originated by the Center and now serving as a national depository for behavioral objectives and related test items.

The CSE Elementary School Test Evaluations, which is a listing of all standardized tests keyed to educational objectives of elementary school education, and evaluated by measurement experts for meaningfulness, examinee appropriateness, administrative usability, and quality of standardization.

The Evaluation Workshop I, designed as a training device to provide school administrators and project directors with an understanding of the kinds of information an evaluation can provide for educational decision-making, and the general procedures and problems involved in selecting, collecting, and analyzing that information.

Each of the above products, as well as others, will be discussed in further detail later in this report. In addition to these products, the Center has also developed a number of tests and other measurement devices for important educational dimensions for which no adequate measures presently exist.

A FRAMEWORK TO GUIDE EVALUATION PRODUCT DEVELOPMENT

In the past year, there has been increasing evidence of a developing consensus on a broader, more comprehensive definition of evaluation. It is the view at the Center for the Study of Evaluation that the judgments made by evaluators must be of practical use to decision-makers in selecting among various courses of action. This view of evaluation also acknowledges the uniqueness of specific situations or programs and the necessity of recognizing that uniqueness in the evaluations as well as in the manner in which the evaluation information is ultimately reported.

With such an emphasis upon the needs of decision-makers and the unique characteristics of specific programs or situations, the Center has formulated a definition of evaluation which focuses on the necessity for meeting the information needs of decision-makers within their own specific situations. Evaluation has thus been defined as the process of *ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.*

This seemingly complicated definition means, in simple terms, that it is the function of the evaluator to provide information to decision-makers that can be used effectively to make decisions about alternative courses of action. For example, in *ascertaining the decision areas of concern*, the decision-maker determines the nature of the domain to be examined, while the evaluator serves to point out inconsistencies, difficulties, or additional data that might modify the decision-maker's views on the relevance of certain outcomes. In *selecting appropriate information* in light of the decision areas to be considered, the task of the evaluator includes specifying the evaluation design of the project and the selection and/or development of instruments designed to provide the information appropriate to the decision areas. In *collecting and analyzing* that information the evaluator must con-

sider various problems associated with the task, depending upon the unit being evaluated, the nature of the decision-maker, and other considerations. Finally, when *reporting summary data* to the decision-maker, the function of the evaluator is to make judgments about the relative worth of the various courses of action.

This general view of evaluation has led the Center to consider the total instructional improvement process in order to identify the normal sequence of educational decisions that might occur. The identification of the educational sequence (or of *one possible* educational sequence) provides insight concerning the kinds of judgments that decision-makers must make and the associated evaluation information needs of these decision-makers.

As can be seen in the chart, the Center has identified five major decision areas of concern and the five kinds of evaluation which provide information for these areas. The decision areas are (1) selection of the appropriate problem or objective to be served (problem selection), (2) selection and design of the program to be introduced which best fulfills the objective (program selection), (3) placing the program in operation in the manner in which it was described (program operationalization), (4) modification of the program in terms of field constraints (program improvement), and (5) verifying the appropriateness of the program for introduction elsewhere (program certification). The five kinds of evaluation are intended to provide information related to these decision areas:

Needs Assessment

It is a reasonable assumption that the decision-maker may be forced to make choices related to school or

Chart 1

Decision Areas and Evaluation Requirements

Evaluation	*Decision Area*
Needs Assessment	Problem Selection
Program Planning	Program Selection
Implementation Evaluation (Intervention)	Program Operationalization
Progress Evaluation (Intervention)	Program Improvement
Outcome Evaluation (Non-Intervention)	Program Certification

program priorities. Given the scarcity of resources within his current situation, a decision-maker might want to make a decision about the problem area most in need of attention. This "problem selection decision" is based in large part on a *needs assessment.* Needs assessment attempts to examine the gap between specific goals and the existing situation. The evaluative problem is essentially one of assessing the needs of students, of the community, and of society in relation to the current status

This Center is one of a system of eight Educational Research and Development Centers funded under the Cooperative Research Act (as amended by Title IV of the Elementary and Secondary Education Act of 1965). The Program was organized as one response to an increased national awareness of the importance of finding solutions to critical educational problems.

More specifically, the R & D Centers program was devised to fill a unique role in relation to other forms of educational research and development, by providing a prime avenue for (a) bringing together a critical mass of interdisciplinary talent and other research resources from the behavioral sciences and other disciplines, (b) focusing on a crucial educational problem area by means of a long-range coordinated attack on large-scale problems, and (c) moving promising innovations through development toward an impact on actual educational practice. Although R & D Centers generally do not carry the innovative process through to final implementation themselves, they are charged with the responsibility for projecting a further route toward that goal by enlisting the interest of a regional educational laboratory, commercial developer, state or local agency, coordinating body, or other appropriate institution.

This Annual Report describes some of the recent accomplishments of one of these centers in its progress towards meaningful education change. The complete list of eight R & D Centers is as follows:

Learning Research and Development Center, University of Pittsburgh (1964)

Center for the Advanced Study of Educational Administration, University of Oregon (1964)

Research and Development Center for Teacher Education, University of Texas at Austin (1965)

Stanford Center for Research and Development in Teaching, Stanford University (1965)

Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin (1964)

Center for the Study of Social Organization of Schools, The Johns Hopkins University (1966)

Center for Research and Development in Higher Education, University of California at Berkeley (1965)

Center for the Study of Evaluation, University of California at Los Angeles (1966)

Also funded through this same program is the *National Program on Early Childhood Education,* which consists of seven University-based centers whose research and development efforts are coordinated through the Central Midwestern Regional Educational Laboratory (CEMREL), St. Ann, Missouri.

* * * * *

The Educational Research and Development Centers are part of a larger set of institutions which contribute in specialized ways to the improvement of educational practice. These include:

The two *Educational Policy Research Centers,* charged with providing a continuing examination of future educational needs and resources for the years 1980-2000.

The two *Vocational Education Research Centers,* established under the provision of the Vocational Education Act of 1963.

The system of 15 *Regional Educational Laboratories,* each of which concentrates on specific problems concerned with the development, demonstration, and dissemination of educational alternatives, materials, and practices for the schools; some of these have close relationships with the Educational Research and Development Centers.

The *Educational Resources Information Center* (ERIC), a nationwide network for acquiring, selecting, abstracting, indexing, storing, retrieving, and disseminating information about educational research and resources, including 20 ERIC Clearinghouses each providing coverage of a particular educational area.

or accomplishments of the system.

Once he has been provided with this information, the decision-maker concerned with the instructional improvement process is able to select from among the alternatives those problem areas (objective dimensions) which need attention or modification.

Program Planning Evaluation

A second decision area of concern deals with program selection. The decision-maker, having made a decision about the specific problem area (or objective) in which his system is deficient, must choose among the existing programs directed toward the achievement of that objective and various alternative programs that might be introduced. These choices are referred to as "program selection decisions." These decisions are in large part based upon information provided in a *program planning evaluation.* In this evaluation stage, the evaluator is generally asked to provide information on the possible future impact of the introduction of several alternate programs. The evaluator's task is to provide, prior to a program's inception, information concerning the potential success of that program.

When such information has been provided, the decision-maker must make a judgment as to which program offers the greatest probability of success within his system. The decision is not completely dictated by the data provided by the evaluator. Even though the evaluator should attempt, within his study, to be as cognizant as possible of the political and contextual variables which bound or restrict the nature of the ultimate decision, it is not likely that he can become aware of them all. Thus, the program actually selected may differ somewhat from the one which looks best in theory.

Implementation Evaluation

Following the decision to adopt a program and at the beginning of the implementation of the program, the decision-maker must be able to determine the extent to which the program has been put into operation in the manner in which it was intended. This "program operationalization decision" is based upon an implementation evaluation.

In an *implementation evaluation* the concern is with providing information on the manner in which the program has been implemented. One relevant question is the degree to which that program, as described in the program selection decision, has been introduced. That is, has the program been introduced and is it operating in the pedagogical manner in which it was intended to be used? Furthermore, the program selection decision, which selected the indicated program, was based on various assumptions about the nature of the student population to be served. If these assumptions were incorrect or are no longer appropriate, then this information, as part of an implementation evaluation, should be meaningful to the decision-maker in determining whether it is appropriate to allow the program to continue.

Progress Evaluation

Having made the decision to introduce a specific program, the decision-maker must be able to make program modifications, as the situation demands, throughout the course of the introduction of that program. These "program improvement decisions" require evaluation information of a certain type. The Center refers to this information as *progress evaluation.*

In a *progress evaluation* the evaluator is called upon to provide data on how the program is functioning. In terms of the short-range objectives of the system, such as the intended objective to be achieved at the end of a specific unit of study, what has been the performance of the student group? Are there noticeable unanticipated outcomes, not a part of the original objective for the program but which, nevertheless, ought to be noted as important information potentially valuable to a decision-maker in making decisions about the program?

The decisions related to progress evaluation are likely to take the form of program changes during the course of the program rather than be decisions made at the conclusion of the program. The evaluative function identified in progress evaluation is designed to be interventionist in nature. That is, during this stage the evaluator is envisaged as one who is concerned with providing information regarding the potential modification and improvement of programs during the process of their introduction. This is directly opposed to the more passive role that might be associated with a researcher-observer who, through his desire to draw valid generalizable conclusions, is careful not to intercede in the process.

Outcome Evaluation

At some point after a program has been introduced into the system and has been properly implemented and modified in line with whatever difficulites have been noted, the decision-maker may wish to consider the potential generalizability of the program. A decision related to the potential generalizability of a program to other educational systems is viewed as a "program certification decision." The evaluation associated with such decisions is referred to as *outcome evaluation.*

In an outcome evaluation the role of the evaluator is modified from the interventionist stance previously described. In order to maintain the generalizability of the situation, it is imperative that the evaluator not be actively involved in the program, and that he attempt to insure that drastic program modifications are not being made concurrently with the evaluation. Most protocols of experimental research would be applicable to such situations.

* * * * *

In considering the situations in which evaluations might occur, the Center has distinguished between the evaluation of educational systems and the evaluation of instructional programs. In terms of the conceptual framework discussed above, the evaluation of educational systems is seen as involving the first two evaluative stages and the evaluation of instructional programs as involving primarily the last three stages.

Evaluation of an educational system requires determining educational needs in terms of the most appropriate objectives for the given system and providing information on the progress of the system relative to these dimensions. When decisions have been made concerning which of the system's objectives are inadequately met, the decision-maker might be concerned with the selection of programs to meet these objectives and the evaluator should provide information relative to the possible impact of various courses of action or programs. At the stage where the program has been selected, the task shifts from evaluating the total system to an evaluation of a particular instructional program.

Evaluation of an instructional program assumes not only the prior assessment of the program or of a larger system, but also that decisions have already been made concerning objectives to be attained and that programs to meet these objectives have been selected. The evaluation of an instructional program, then, normally begins after the decisions relating to needs assessment and program planning have been made. The focus, therefore, is primarily upon the last three stages of evaluation. Here the evaluator is concerned with providing evaluative information on program implementation, progress, and outcome. The evaluator should provide the decision-maker with information concerning the extent to which the program has been implemented as it was described in the program plan, the extent to which the program is proceeding satisfactorily and, finally, information concerning its potential generalizability to other situations.

The conception of evaluation discussed above is reflected in the Center's organizational structure and the goals of its current activities, which are contained within three programs: (1) the Program on Evaluation of Educational Systems; (2) the Program on Evaluation of Instructional Programs; and (3) the Program on Evaluation Methodology and Theory. At this time, the Center supports two major projects in the first program, one in the second program, and one in the third program.

THE PROGRAMS

EVALUATION OF EDUCATIONAL SYSTEMS

The objective of this program is to construct procedures and methodology for evaluating educational systems. This program is primarily concerned with the needs assessment and program planning stages of evaluation. It is comprised of two projects: (1) *The School Evaluation Project* and (2) *The Higher Education Evaluation Project.*

EVALUATION OF INSTRUCTIONAL PROGRAMS

The purpose of this program is to construct procedures and methodology for evaluating instructional programs. Program evaluation is primarily concerned with the implementation and outcome stages of evaluation. At the present time, one project, the *Project for Research on Objective-Based Evaluation (PROBE),* is supported under this program.

EVALUATION METHODOLOGY AND THEORY

The objectives of this program are to develop measurement procedures appropriate to many kinds of evaluation settings, to work towards the development of evaluation theory, and to engage in other activities that assist governmental agencies. The latter objective is engaged in only to the extent that time and financial resources permit. Of primary importance in the program is the *Training Materials Development Project* which is currently developing simulated evaluation materials for use by evaluators and school administrative personnel. This project permits CSE to sharpen its views on evaluation theory. The Center also engages in *Opportunity Projects* in which for minimum financial costs, CSE can produce meaningful products or research results.

The following Program and Project Register lists the principal investigators for each of these projects.

PROGRAM AND PROJECT REGISTER

Center for the Study of Evaluation

UCLA

Code Number	Title	Principal Investigator
01	EVALUATION OF INSTRUCTIONAL PROGRAMS	Rodney Skager
0106	Project for Research on Objective-Based Evaluation	Rodney Skager
02	EVALUATION OF EDUCATIONAL SYSTEMS	Marvin C. Alkin
0201	School Evaluation Project	Ralph Hoepfner
0202	Higher Education Evaluation Project	C. Robert Pace
03	EVALUATION METHODOLOGY AND THEORY	Marvin C. Alkin
0305	Training Materials Development Project	Stephen Klein
0311	Opportunity Projects	Richard Seligman

PROGRAM ON EVALUATION OF EDUCATIONAL SYSTEMS

SCHOOL EVALUATION PROJECT

The School Evaluation Project is designed to develop and field test sets of procedures which might be employed by school evaluators and administrators engaged in evaluating schools – preschool, elementary, and secondary. The project is attempting to capitalize upon the state of current knowledge to develop evaluation procedures which are appropriate especially to the first two evaluation stages of the Center model and to education at various levels. The Center is concerned with developing procedures which will enable school principals and others to use information effectively in making valid decisions for improving student performance. The project is currently field testing an Evaluation KIT which is made up of a series of booklets describing how to conduct a needs assessment of an elementary school's student output.

The KIT is designed to be used primarily by the elementary school principal to aid him in selecting, collecting, and interpreting the information he will need in making educational policy decisions. The KIT is further designed so that, with minor alterations, it can be extended to educational problems at other levels. The KIT, essentially limited to the first stage, needs assessment, in CSE's conception of evaluation, will be extended from the needs assessment stage to the program planning stage of the CSE model. The KIT presently builds to the program planning stage and leads to questions and problems best treated in a "how-to" fashion for that stage of evaluation.

Current plans are developing toward the transposition of the KIT from the elementary school level to other educational levels. Initial concentration would be for Preschool-Kindergarten, and then for Junior and Senior High School. Other levels will be attacked by studying anew the goals for each level of education, categorizing them meaningfully, producing needs assessment packages appropriate to the new categories, and evaluating appropriate tests according to the categories.

During the 1969-1970 fiscal year, the Elementary School Evaluation KIT has been developed, field tested within California, and refined for further field testing. Within a year the present KIT will have been extensively and intensively field tested and will be ready for dissemination. Development of the KIT has also led to several by-products as results of pursuits closely parallel to the ongoing research and development. The Project's accomplishments are described below.

1. *Elementary School Evaluation KIT: Needs Assessment*

Completed. Four booklets of the KIT have been rewritten and reorganized, based on results of initial field testing. These four booklets are entitled:

I. Description – (organization of KIT and suggestions for its usage)

II. Choosing Goals – (how to determine priorities of felt needs regarding student performance)

III. Selecting Tests – (how to select tests that will yield the needed kinds of information)

IV. Collecting and Analyzing Information – (how to collect the needed information, interpret it, and communicate its implications)

These first four booklets of the KIT have been extensively field tested in the state through the cooperation of the California Elementary School Administrator's Association (CESAA), headed by Dr. Edward Beaubier. The CESAA principals have been chosen to represent the state geographically.

The fifth and last booklet of the KIT, Selecting Critical Need Areas, is concerned with delimiting goal priorities for program implementation through consideration of utilities of programs for areas, probabilities of their payoffs, and estimates of their costs.

Planned. The fifth booklet of the KIT, expected to be completed by fall, 1970, will be field tested with the CESAA principals and superintendents who have already worked with booklets I through IV. On the basis of the feedback from the field testing, the KIT will be printed in smaller booklets for continued field testing. The planned field testing will be of two types – case-study and nation wide.

In the intensive case-study approach, the KIT will be implemented in two or three selected schools for study of problems of implementation. This will be done through intensive interviewing techniques with a CSE staff member making regular visits to the schools and working with their staffs. Without a school's total commitment to following through with the KIT, the real and dynamic problems, uncoverable by the case method, will not emerge. Therefore, the case-study schools will be selected most carefully. Further extensive field testing with larger samples have also been planned. This plan would involve having interested elementary school districts or systems throughout the country prepare proposals for KIT implementation, specifying the users, the meetings planned, and the extent and direction of implementation. The Project would require that each participating system complete detailed and structured periodic status reports, which would be designed to maximize pertinent feedback on the KIT's implementation, and would be further utilized in the production of the final form of the KIT.

Upon completion of the field testings of the KIT, by about mid-1971, a final edition will be produced and a delivery system will be implemented. Alternatives to the delivery system include CSE's publishing of the KIT (with continued feedback study at a marginal level), or having a commercial publisher take over the task. In either case, the Center is paying attention to considering the form and presentation of the material in order for it to be usable by the prospective audience.

2. *Elementary School Evaluation KIT: Program Planning*

Planned. Initial plans for the second KIT in the elementary series are briefly outlined below:

I. Implementing Evaluation Decisions – (a directory of potential problem areas, their alternative solutions, and getting such decisions to work in the school)

II. Finding Information on Programs – (a review of sources of general and specific program information)

3. *Preschool-Kindergarten School Evaluation KIT: Needs Assessment*

Completed. A compendium of goals for early education, based on a comprehensive search of the stated objectives of programs and theories, has been developed and the evaluations of tests and measures has begun. Minor modifications of the Elementary School procedures will eventuate in another major CSE product. The Center is convinced that publication of this work will have major impact upon the measures used in evaluating early childhood education.

Planned. Based on experience with the Elementary KIT, a parallel form of the test evaluation book for early education will be available in late 1970.

4. *Attitude-Toward-School Instrument*

Completed. Based on the analysis of needs preparatory to the writing of Booklet II of the Elementary School Evaluation-Needs Assessment KIT, it became obvious that favorable school-attitude was a high priority student goal, but one for which there is no adequate published assessment instrument. The School Evaluation Project initiated the construction of a non-reading instrument to assess eight dimensions of school attitude, and to fit well with the requirements for effective test instruments as defined by the CSE test evaluation system.

The Attitude Toward School Questionnaire (ASQ) was constructed along these lines, administered to several hundred local first-grade students, intensively analyzed, and discussed in a CSE Report entitled "Development of a School Attitude Questionnaire for Young Children."

5. *CSE Elementary School Test Evaluations*

Completed. A bound book detailing the procedures and results of the application of the MEAN criteria to all available published tests appropriate for the elementary school has been published by UCLA. This publication was necessitated by the overwhelming positive response of educators and the consequent need to release the information as soon as possible, rather than delay it until

publication of the whole KIT.

The procedure utilized in the development of the *CSE Elementary School Test Evaluations* is, basically: (1) a listing of most of the major goals of elementary school education was developed, (2) all published standardized tests were categorized in terms of which objective they most closely met, (3) a quantifiable test rating (MEAN) was developed and, finally, (4) each test or subtest was evaluated using the MEAN rating system in terms of the appropriateness of the test at four grade levels (1, 3, 5, and 6). Each test was, therefore, evaluated in terms of a quantitative rating of its *M*easurement validity, *E*xaminee appropriateness, *A*dministrative usability, and *N*ormed technical excellence.

The Center feels that the *CSE Elementary School Test Evaluations* serve a most useful purpose and provide information not currently available. CSE views its major advantages as (1) conciseness, which is not offered by any other test reviews, thereby seriously diminishing their use for most educators; (2) currency, which again is not provided in current compendiums, but found only in scattered test reviews; (3) educational relevance, since all tests are evaluated against consensual educational objectives; (4) objectivity, since the Center has no formal or informal relationships with test publishers; and (5) consistency, as a single set of standards against which all tests are evaluated is employed.

The Center is convinced that the publication of the *CSE Elementary School Test Evaluations* will have major impact upon the tests presently selected in evaluations, and in improving the quality of tests available in the next several years. This, the Center feels, is a significant contribution to American education.

6. *Technical Reports on Early Childhood Tests*

Completed. Another outcome of the MEAN test evaluation was the critical consideration of several aspects of tests available for young children. This critical attitude resulted in two technical reports: *The Language of Tests for Young Children*, and *The Unintentional Memory Load in Tests for Young Children.*

Among the findings of *The Language of Tests for Young Children* was the influence of linguistic variables on children's test performance, and that the results of tests of aptitude and achievement with young children are often influenced in ways not directly related to the abilities being tested. The results of the *Unintentional Memory Load in Tests for Young Children* show that memory processes are essential to the comprehension of language and that memory processes are of sufficient importance to be measured properly and not confounded with other skills. The Center feels that the findings of these studies are likely to produce an impact on the future construction of tests written for young children.

7. *Project's Catalytic Relationship to Commercial Test Publishers*

Completed. Given the options of "sitting on" important evaluative information regarding currently available tests, coldly and remotely criticizing publishers, or working with the publishers for the improvement of their instruments, the Project has opted for the last alternative. This program is widened by the following activities: immediate distribution to test publishers of the relevant technical reports and prototypes of test instruments developed by the Project, and a planned conference to which all test publishers will be invited. The conference, to be held in the summer of 1970, will concentrate on (1) typical instrument shortcomings as revealed by the MEAN evaluation system and how improvements can be implemented, and (2) goals areas devoid of measurement devices and how these voids can be filled well and profitably.

8. *Science Test Prototype*

Completed. From the MEAN evaluations, major instrument shortcomings became obvious in the goal areas of science. The Project has therefore developed a prototype science test which concentrates its assessment on the higher-level aspects of science education. Initial contacts have been made with a well-known test publisher to pursue further development and standardization of the prototype instrument.

9. *Hierarchical Goals Charts*

Planned. In order to link the goals utilized by the School Evaluation Project in its needs assessment and test evaluation procedures to the instructional objectives of the PROBE variety, the Project is working jointly with the curriculum staff of a Southern California school district (Newport-Mesa) in the construction of goals charts which will analyze the generalized goals successively into curriculum objectives and then instructional objectives. Response of CSE visitors to the idea of the hierarchy of goals has been so positive that plans are now underway for the printing of charts of the objectives to aid educators in locating and finding objectives at all levels.

10. *Atlas of Scores*

Completed. The Atlas of scores, projecting standardized test scores for various types of schools which are categorized by such variables as geographic area, racial composition, etc., has been completed for the Coleman tests. Since the Project could not atlas all, or even the best tests, it was decided to adopt the policy of goading test publishers into offering such a norming in their manuals. One way of doing this would, of course, be to allow the cooperating publishers to exploit our Atlas tables as much as they wish. Competition may then bring about upgrading.

Project Professional Staff

Paul Bradley
William Doherty
James Dyer
Ralph Hoepfner,
Project Director
Marvin Hoffenberg
Patricia Jansen
Guy Strickland
Louise Tyler
Dale Woolley

* * * * *

HIGHER EDUCATION EVALUATION PROJECT

The Center believes that certain characteristics of higher education differ in many respects from those of other educational levels and thus present a unique setting for the study of evaluation in education. Higher education is neither as homogeneous nor as structured as elementary or secon-

dary education. It is characterized by a diffusion and decentralization of authority and influence and by a diversity of goals, objectives, and programs within and among colleges and universities. Furthermore, as the highest level in the formal education process, the evaluation of the outcomes of this educational experience is of special social importance; it provides an opportunity to understand the role that the most educated segment of the nation's populace plays in our society. Thus, unlike other levels of education, the evaluation of higher education includes the development of criterion measures that deal with behavior beyond formal schooling.

Some of these characteristics of higher education pose unique methodological problems for evaluative research. For example, granted that all complex programs consist of many different activities, and that their total number and variety are too large to manage or even to investigate, one needs both to develop guidelines for estimating how much to include and, through various data reduction methods, to reduce the set to more feasibly managed dimensions without losing the overall perspective or any potentially important uniqueness. Furthermore, if one grants that all programs have multiple consequences (many of which may not be objectives or intentions of the program), then it is not sufficient only to base one's evaluation around the question, "what are the objectives?" Therefore, suggestions or rules need to be devised for determining the range of outcome or consequences that ought to be included. CSE feels that reasonable solutions to these problems would contribute to a more powerful definition of evaluation and a more productive practice of evaluation.

Objectives of the Higher Education Evaluation Project

The objectives of the Higher Education Evaluation Project are (a) to provide better instrumentation of potential use for needs assessment evaluation of higher education, (b) to demonstrate the importance of broad evaluation for wise decisions and judgments, (c) to create models or examples of higher education for national visibility and potential influence in policy decisions and for local use as a part of the program planning stage evaluation, and (d) to examine those evaluation problems which are peculiar to higher education, and to explore the implications that these problems present in considering a general model for evaluation in education.

1. *Higher Education Survey*

The major activity the Project embarked upon to meet the above stated objectives was a national survey of 90 colleges and universities. The survey has served several purposes: (1) it has provided a field test for a variety of scales and indices that can subsequently be used in institutional self-studies; (2) the survey is in itself an evaluation of higher education and as such will enable one to view higher education from different perspectives in relation to different objectives and outcomes, in relation to the diversity of student characteristics and backgrounds, in relation to a variety of college experiences, in relation to the effects of different educational programs, and in relation to the particular settings or institutional contexts in which the college experience has occurred. Thus, the survey will provide a national perspective within which different colleges and universities can view themselves. Such information will be necessary for the second stage of the CSE model for evaluation, program planning; and (3) the survey also provides an opportunity to explore the consequences of different modes of data analysis and subsequently should have an influence on the practice of evaluation.

Completed. Three basic questionnaires were sent to samples of freshmen, upperclassmen, and alumni representing 90 institutions. Samples of both freshmen and upperclassmen were obtained in 75 of these institutions, and samples of all three populations, in 61 institutions. Data collection was completed in February, 1970. The data have been put on magnetic tape and checked for possible errors. Data analysis has begun on the scales and indices contained in the questionnaires.

Planned. This fall, CSE will direct its efforts toward examining the data with regard to the second and third purposes of the survey outlined above.

2. *The Evaluation Kit*

One of the major thrusts of the CSE Higher Education Evaluation Project for FY 69-70 has been toward the development of a flexible evaluation kit for institutional self-evaluation. The evaluation kit will contain a variety of short scales and indices to measure a broad range of criterion and contextual variables that can be used as separate evaluation instruments. In this way, institutions can select from the set those measures which are most relevant for their own purposes. Insofar as normative data can be provided for the measures, institutions can compare themselves to other colleges and universities and, more specifically, to institutions similar to their type. A manual will accompany the kit to guide the user in the selection and administration of instruments, the analysis of the results, and the effective interpretation and use of the information.

The scales and indices that will comprise the evaluation kit reflect the efforts of the Project to extend the range of criterion and contextual variables that might be used in national or local evaluations of higher education and to insure that such measures are feasible to use. Most of the instruments currently available in the field are related to intellectual performance and attainment or to the assessment of personality traits and attitudes. Even if institutional personnel were able to administer these measures in less than several hours, they would not have obtained the type of information necessary to evaluate a diverse system serving many kinds of objectives. Interests, activities, values, beliefs, and many other kinds of criteria are also relevant to the purpose of higher education. Thus, an effort was made by the Project to develop short, reliable measures that would have the potential for reflecting a wide range of outcomes and a range of possible influences related to these outcomes.

Planned.

A list of the potential measurements which will comprise the kit follows, categorized by the issue to which they most pertain:

(a) *Governance:* Institutional Stance Toward Student Freedom and Discipline; Measurement of Management Styles; Campus Morale.

(b) *The Learning Process:* Short-Form Verbal Aptitude Test; Student Learning Styles; Faculty Teaching Orientation; Student Perception of Educational Objectives; Staff Perception of Educational al Objectives; Perception of College Experience; Satisfaction with College Experience.

(c) *General Environment:* Curricula Environment Assessment; Short Form College and University Environment Scales; Staff Characterization of the Institutional Environment.

(d) *Institutional Innovativeness:* Organization Adaptiveness; Institutional Readiness for Innovation.

(e) *Attitudes, Values and Behavior:* Involvement in Campus Life; Personal Traits and Intellectual Interests; Alienation and Belongingness; Peer Group Patterns; Measurement of Political, Cultural, Scientific, Religious and Community Interests and Involvement; Reading Habits; Activism; Social Awareness (including sub-scores for knowledge and/or attitude about the Changing Society and Environment, the Role of Government, the Role of Women, Civil Rights, Liberalism, and Social Involvement); Index of Vocational Goals or Status; Index of Occupational Characteristics Anticipated or Realized.

Completed and pending. Most of the above measures were embedded in the national survey described above. Psychometric analysis of the relevant data gathered from this survey is nearing completion. These analyses will serve as bases for making decisions concerning specific item sets for preliminary scales and indices. Other measures have been constructed and field tested to a limited extent. The most useful item combinations for these measures are being identified on the basis of the field test results and will be presented as scales for further field testing this fall in collaboration with the Regional Educational Laboratory for the Carolinas and Virginia (RELCV). (This collaboration was established formally through a joint conference on February 3, 1970, initiated by the Center.) It is anticipated that at least 20 of the scales and indices listed above will comprise the preliminary evaluation kit.

In order to assure that the scope of the evaluation kit will be relevant for a variety of users, the Project staff has been working with advisory groups of administrators, faculty members, and students in the Southern California area, with institutional research officers in a variety of colleges and universities in other parts of the country, and with a consortium of colleges affiliated with the Regional Educational Laboratory for the Carolinas and Virginia.

3. *The Development and Refinement of Strategies and Procedures for Institutional Self-Evaluation*

Essential to the effective implementation of the kits is knowledge obtained from monitoring the use of the evaluation techniques in order to further develop strategies for evaluation procedures and ultimately to aid in the formation of a model of evaluation that is appropriate for higher education.

Completed. The staff has been working with several other groups to obtain their suggestions and reactions to proposed procedures for the implementation and use of the evaluation kits. Meetings have been held for this purpose with the administrative committee and with representatives of several colleges and universities that have participated in the survey.

Planned. RELCV will assist CSE in the consistent field testing of the evaluation kit and to monitor the use of the evaluative techniques in order to assess the conditions under which they contribute to program understanding and improvement. Dr. James Trent, a member of the CSE professional staff, will spend six weeks at the Regional Laboratory this summer to facilitate these collaborated efforts.

4. *Institutional Profiles*

The institutional profiles will be sent to the 90 institutions that participated in the national survey. CSE believes that the evaluations represented by the institutional profiles will be of considerable use to the institutions of higher learning in their decision making. Moreover, the institutional survey will constitute base line data and an initial evaluation system for these institutions.

Completed. The project staff met at least once with representatives from 26 colleges attending various conferences to obtain their opinions on content and format of the institutional profiles which would be of maximum usefulness for their evaluation purposes.

Institutions Meeting With CSE

Albion College
Albion, Michigan

Beloit College
Beloit, Wisconsin

Bridgewater College
Bridgewater, Virginia

University of California,
Berkeley
Berkeley, California

Carleton College
Northfield, Minnesota

Colgate College
Hamilton, New York

University of Colorado
Boulder, Colorado

Concordia Teacher College
River Forest, Illinois

Denison University
Granville, Ohio

Duke University
Durham, North Carolina

Fayetteville State College
Fayetteville, North Carolina

Indiana State University
Terre Haute, Indiana

Macalester College
St. Paul, Minnesota

Memphis State University
Memphis, Tennessee

Michigan State University
East Lansing, Michigan

State University of New York
at Geneseo
Geneseo, New York

Purdue University
West Lafayette, Indiana

Radford College
Radford, Virginia

University of South Carolina
Columbia, South Carolina

Susquehanna University
Selinsgrove, Pennsylvania

University of Tennessee
Knoxville, Tennessee

University of Utah
Salt Lake City, Utah

Virginia Military Institute
Lexington, Virginia

Virginia Polytechnic Institute
Blacksburg, Virginia

Wayne State University
Detroit, Michigan

Westmont College
Santa Barbara, California

Wittenberg University
Springfield, Ohio

The preparation of the institutional profiles has involved the following activities:

(1) Layout of the National Survey data and computer programming and output so that each participating institution can compare the responses of its students and graduates, by sex, to those of the sample at large and those of all institutions of their own type (i.e., General Liberal Arts, Select Liberal Arts, General Colleges and Universities, etc.).

(2) Arrangement of select cross-tabular analyses to accompany the normative data for the sake of a fuller indication of their usefulness in assessments of institutional status, outcome, and key educational variables associated with outcomes.

(3) Preliminary set-up of data for interpretive commentary to accompany the report of the normative data.

(4) Preliminary draft indicating: the relevance of the profiles and normative data to institutional evaluation; the usefulness of additional analyses and further use of evaluation packages available from the Project for additional self-evaluation; and methods for undertaking these evaluations.

It is anticipated that the institutional profiles will be completed by October, 1970.

Project Professional Staff

Marsha Ehrenberg
Michael Huberman
Sonja Jacobsen
Ann Morey,
Executive Officer

C. Robert Pace,
Project Director
James Trent

* * * * *

PROGRAM ON EVALUATION OF INSTRUCTIONAL PROGRAMS

PROJECT FOR RESEARCH ON OBJECTIVE-BASED EVALUATION (PROBE)

The Center believes that the most basic need in the evaluation of specific instructional programs is the proper identification of well-stated objectives and of test items to measure these objectives. Thus, while it is recognized that there are other elements to be considered in evaluating an instructional program (particularly in implementation evaluation), the Center has concentrated its resources within the project on attempting to make a major breakthrough in the area of objectives and related items. CSE feels that it has already made significant contributions to the area of objective-based evaluation and that work in progress offers even greater prospects.

PROBE has been engaged in accomplishing a number of complementary tasks designed to lead to major improvement in the way in which instructional programs are evaluated using objective-based measures. These activities are (1) to develop objectives and related test items in those subject matters most commonly taught, (2) to devise procedures for obtaining quality control of objectives and items, (3) to facilitate the implementation and use of objective and item collections, and (4) to develop sophisticated systems (perhaps computer based) to allow users to generate evaluation systems for multiple purposes and at various levels of objective hierarchical categorization (e.g., the instructional unit behavioral objective versus the end of semester schoolwide objective).

1. *The Instructional Objectives Exchange*

Completed. One of the major components of PROBE was the Instructional Objectives Exchange (IOX), the function of which was to provide for the depositing, developing, and disseminating of measurable objectives and items for use by educators. IOX was founded on the premise that the majority of teachers were either not well enough prepared to engage in the task of generating measurable instructional objectives which could be used as a basis for planning and evaluating instructional activities, or were too busy to perform this function well. It did seem possible, however, for teachers to engage in the selection of measurable objectives from collections of these objectives made available through an agency such as IOX. This would not only permit local autonomy in the selection of goals, but also facilitate the organization of instruction around statements of measurable objectives.

In order to develop collections of objectives in a number of subject fields for both primary and secondary schools, objectives and related items were collected from school districts all over the country. These objectives and items had been contained in curricular material submitted by hundreds of school districts. They were compiled, revised, and rewritten to conform to standards established by PROBE. In cases where objectives and items did not exist, they were developed by members of the PROBE staff. As a result of this concerted effort PROBE, through the IOX activity, published *sixteen* collections of objectives, in various subject areas and grade levels. The collections already available are:

MATHEMATICS, K-3 – This collection emphasizes the introduction of concepts and skills. The content area includes sets; numbers, numerals and numeration systems; operations and their properties; measurement; geometry, relations, functions, and graphs; probability and statistics; applications and problem solving; and mathematical sentences, order, and logic. Four sample items are listed for each objective. (275 objectives)

MATHEMATICS, 4-6 – This collection emphasizes intermediary concepts and skills. The content area includes sets; numbers, numerals, and numeration systems; operations and their properties; measurement; geometry, relations, functions, and graphs; probability and statistics, applications and problem solving; and mathematical sentences, order, and logic. (246 objectives)

MATHEMATICS, 7-9 – This collection emphasizes those concepts and skills structural to the discipline of math. The content area includes sets; numbers, numerals, and numeration systems; operations and their properties; measurement; geometry, relations, functions, and graphs; probability and statistics; applications and problem solving; and mathematical sentences, order, and logic. (265 objectives)

READING, K-3 – This collection emphasizes word recognition, comprehension, and study skills. Each of these areas is respectively subdivided into the major categories of phonetic and structural analysis, literal, interpretive, critical and vocabulary comprehension, and the work skills of alphabetizing, reading rate, organization, and use of references, pictorial and graphic material. (approximately 314 objectives)

READING, 4-6 – This collection emphasizes the areas of word recognition, comprehension and study skills. Each of these areas provides objectives for the sub-categories to develop and extend the skills acquired in grades kindergarten through three. (approximately 183 objectives)

READING, 7-12 – This collection emphasizes structural analysis, critical comprehension, and study skills. Objectives for selected phonetic analysis skills are also included, as well as extensive coverage of the major sub-categories. (approximately 193 objectives)

ENGLISH LITERATURE, 10-12 – This collection is designed to develop the students' ability to analyze literature and to evaluate its effects. Content areas include poetry, the novel and drama. An average of six sample items are listed for each objective. (25 objectives)

LANGUAGE ARTS, K-3 – Contents in this collection include writing skills, simple paragraph and letter forms; grammar skills, listening and speaking skills. (approximately 40 objectives)

LANGUAGE ARTS, 4-6 – Contents in this collection include structure and types of sentences, parts of speech, capitalization, punctuation, linguistics (word analysis), composition, literature. (188 objectives)

LANGUAGE ARTS, 7-9 – Contents in this collection include reference skills, listening and speaking skills, composition, literature. (96 objectives)

AUTO MECHANICS, 10-12 – This collection reflects major behavioral objectives required in a comprehensive course in automotive tune-up and repair. (140 objectives)

SOCIAL SCIENCE (Geography), K-9 – This collection reflects major social science concepts in the discipline of geography. (158 objectives)

BIOLOGY, 10-12 – This collection emphasizes processes of inquiry and laboratory work, either directly or indirectly. Although this sequence is based on the Biological Sciences Curriculum Study (BSCS), the objectives are designed to make them serviceable to any program. (approximately 62 objectives)

HEALTH (Nutrition), K-6 – This collection reflects the major concepts related to Nutrition – Man and his Food. Content area includes daily food choices using the four food groups, nutrients from food, how the body uses food, food processing, consumer education in advertising and merchandising, and cultural and social uses of food in man's environment. (approximately 30 objectives)

MUSIC, K-6 – This collection reflects major concepts, fundamentals and applications in music appreciation. (approximately 50 objectives)

PHYSICAL EDUCATION, K-3 – The content area includes perceptual motor, sensory motor, locomotor skills, non-locomotor skills, balance, eye-foot skills, eye-hand skills, and dance. (approximately 50 objectives)

The Center has participated in the development of an additional set of collections geared to secondary education. The cost of developing these collections was in large part supported by a grant from the Educational Systems for the Seventies (ES 70) program. The collections will be available in the Fall of 1970 and include:

- English Skills, 7-9
- English Skills, 10-12
- English Literature, 7-9
- English Literature, 10-12
- English Grammar, 7-12
- Business Law, High School
- General Business, High School
- Secretarial Skills, High School
- Bookkeeping, High School
- Home Economics, Junior High
- Home Economics, Senior High
- General Metals, High School
- Electronics, High School
- Mechanical Drawing, High School
- Woodworking, High School
- Spanish
- General Math, 10-12
- Advanced Math, 10-12
- American History

To this date, nearly 25,000 objective-item collections have been distributed to educators in every state of the union, and beyond, and the comments received have been enthusiastic. There is every reason to believe that the Center has made a major impact on educational practice, and that PROBE efforts have provided educators with much needed tools to improve objective-based evaluation in the schools.

Since it is CSE's view that it cannot accomplish everything with its limited resources, the Center has had to make basic decisions about when an activity is no longer research and development and more properly a service activity. This is the case with the Instructional Objectives Exchange. In CSE's view, and in that of the U.S. Office of Education, the Center had accomplished its purpose with IOX. That is, it had been demonstrated that objective and item collections (at the behavioral objective level) could be created with cooperation from school districts, states, and others, and could be made available to the field. While the collections are not complete nor comprehensive in coverage, they are, nonetheless, considerably better than what is generally available. The task of continuing to refine objective and item collections of the IOX type has been taken on by the Instructional Objectives Exchange, Inc., a nonprofit California corporation. The new IOX, Inc. will work closely with CSE as the Center pursues its further objectives of (1) developing procedures for implementing objective-item collections of the IOX type into schools, and (2) developing more sophisticated systems of objective classification in order to provide for objective-based evaluation needs of different types.

2. *Quality Control and Implementation Efforts Completed.* PROBE, however, is committed to improving both the quality of the objectives which have been included in its collections, and the methods by which these objectives are implemented in the schools.

Although the majority of the objectives contained in PROBE Collections were based upon goals and aims contained in existing curriculum material, the objectives and items developed for inclusion in the collections had never been tried out in school systems in their present form. Several methods were developed to insure that these objectives would be of practical value to educators. One was to expose the objectives and items developed by PROBE to reaction groups composed of well experienced teachers *before* they were put into collections for nationwide distribution. Another method was to indicate, within each publication, that the objectives and items contained therein might not have been perfected, in spite of the careful perusal by reaction groups, and that, in order to revise the objectives, valuable information might be gathered from the field. To this end, each collection contained a questionnaire designed to elicit information from the user about the content and the form of objectives and items contained in the collections.

The first basis for revision, comments of reaction groups, was designed to create as good a product as possible *before* this product was actually tried out in the field; the second basis for revision, the questionnaire, was designed to make corrections in objectives and items based on the results of these trials.

Moreover, as indicated above, PROBE was concerned with perfecting methods of implementing objective-based systems for evaluating programs. In order to provide a

constant and reliable source for objective and item revision, and in order to establish a means by which a system of implementation could be developed, it was deemed necessary to create a national association of pilot schools. The National Network of Schools was formed for this purpose and to monitor the quality of the revisions suggested by the reaction groups and the actual users of the collections.

Member districts in the Network.

Cherry Creek School District
Englewood, Colorado

Cincinnati Public Schools
Cincinnati, Ohio

Edina Public Schools
Edina, Minnesota

State of Hawaii
Department of Education
Honolulu, Hawaii

Lake Washington School District
Kirkland, Washington

Los Alamitos School District
Los Alamitos, California

Los Angeles Unified School District
Los Angeles, California

Manchester School District
Manchester, New Hampshire

Palos Verdes Peninsula Unified School District
Rolling Hills, California

Point Pleasant Beach Public Schools
Point Pleasant Beach, New Jersey

It was agreed at the outset that two criteria would be used to evaluate the mutual effectiveness of this association between PROBE and the Network of Schools. The first was the extent to which the Network succeeded in fulfilling the purposes for which it had been formed; the second was the extent to which PROBE proved useful to the individual schools which comprised the Network. In terms of the first criterion, the Network not only succeeded in critiquing the collections developed by PROBE, but it identified, as well, gaps in the existing materials and, under PROBE staff supervision, developed additional objectives and sample items to fill these gaps. Some schools have succeeded in developing a complete sequence of objectives and related items in specific subject areas for use in the classroom on the basis of PROBE products. Moreover, member schools have begun to develop objective-based evaluation utilization systems suited to their own particular needs. The PROBE staff is convinced that the Network has succeeded beyond PROBE's greatest expectations in accomplishing the ends for which the Network was designed.

The method employed for evaluating the effectiveness of PROBE for the Network schools is based mainly on comments made by the staff of these schools concerning PROBE's effectiveness in facilitating their educational efforts. One source is the use of questionnaires to document the reactions of school staffs. In all cases where questionnaires were used, 95% or more of the personnel testified that PROBE was of great use to them. Another source of evaluation is school publications. The following quotation is taken from a recent publication of one of the Network schools: "Beginning with the second semester of the 1969-70 school year, the Project for Research on Objective-Based Evaluation (PROBE) had an immediate impact on teachers, materials, and students in the Edina Schools. Edina designated the mathematics department of Valley View Junior High School and the teachers of mathematics in the Creek Valley Elementary School to be the pilot schools of this project. To broaden the influence, Cahill Elementary School was named a companion school, and one teacher from Countryside Elementary School became an observer. These teachers used the Instructional Objectives Exchange collection to select objectives, prepare pretests, and organize learning activities. During the first month about 1,100 students were affected by new methods of instruction designed to meet individual needs."

A third source is letters received by PROBE from the member schools. The following is a quotation from one of these letters. "We need more help from the Network and from all you great people at UCLA's Center for the Study of Evaluation. It would be very discouraging if the Network could not continue to help bring about very positive changes in the educational climate all over the country. Yes, we must find whatever means possible to make this project a continuing one. We need it in New Hampshire. The Northeast needs it. The Kids need it."

At the last Network meeting, all the schools expressed great desire to maintain the involvement which had been established with PROBE and with the Center. Moreover, the Network felt that so much had been gained by the individual schools, that a National Newsletter should be developed to spread the accomplishments of this small nucleus of educators.

Planned. These efforts to create individual utilization systems have been unified into a generalizable user's guide to implement PROBE objectives of the IOX type in the school. After the Network's summer conference, the guide will be field tested in each Network school during the coming year, revised where necessary, and eventually disseminated to all interested school districts.

3. *Total Evaluation System: Prototype-Reading*

CSE envisages the day when complete objective-based evaluation systems will be available for almost every subject area. These systems will contain objectives hierarchially ordered in such a way that potential users can develop or obtain objective-based tests related to whatever level of objectives they are interested in – be it classroom unit objectives, semester objectives, or schoolwide objectives. CSE believes that the user will be able to select objectives at the level of his interest and obtain computer print-out tests specifically geared to his selected objectives.

The Center is attempting to build one such total evaluation system in the area of reading as a prototype. The activity utilizes as input the three collections of reading objectives and items assembled by IOX and proceeds from that point. The reading area has been selected for the initial PROBE efforts since it is clearly the area in which there is presently the greatest interest nationally in the improvement of instruction.

The PROBE evaluation system for the reading program consists of four major elements: (1) instructional objectives; (2) related test items; (3) a retrieval system and user's guide; and (4) a delivery system for implementing the program. The ultimate objective is to develop a delivery plan for utilizing objective-based evaluation procedures. It is envisaged that there will ultimately be a complete system for providing tests specifically geared to the objectives selected for a specific program in a school district, school, individual classroom, or any other level of aggregation.

The Center believes that objective-based evaluation systems must be flexible enough to provide a variety of uses in terms of content, sequencing, and generality of measurement. Even evaluation systems, however, must stand the

test of evaluation, and PROBE is no exception. The classroom feedback use of PROBE must ultimately stand the test of being directly related to student achievement. Students whose teachers use a classroom feedback system based on PROBE materials should show higher achievement than do otherwise similar groups of students whose teachers do not use PROBE. All instructional devices must be directly tied to desirable changes in students. In this sense the classroom feedback use of PROBE is *interventionist* in the learning process and frequently would contribute to decisions about program improvement as described earlier in this report.

In contrast, the use of PROBE for the macro-evaluation of instruction is *non-interventionist* in terms of the period of time in which the information is collected. If the evaluator is to provide information to be used in making decisions about problem selection, program selection, or program certification as described earlier, then the results obtained must not be in part a function of the evaluation itself. The ultimate criterion for evaluating classroom feedback use of PROBE is that desirable growth occur in students in an ongoing program utilizing PROBE. The ultimate criterion for judging the worth of PROBE in evaluating programs and systems is that improved programs are developed or selected where they are found to be necessary.

PROBE research and development activities based on evaluation systems represent a natural growth from the materials collected and organized under earlier phases of work. In the initial development of a PROBE evaluation system the focus is on a single content area to develop a comprehensive system for assessing instructional goals. Reading has been selected to be the content area due to its great importance in nearly all aspects of education.

The system under development is known as SOBE-R (System for Objective-Based Evaluation — Reading). In this system, the reading domain is divided into six major areas. Each area is subdivided into curricular components in a tree-like, branching fashion. Ultimately, behavioral objectives are written for the smallest curricular subdivisions. That is, the curriculum is defined in terms of behaviors expected to be acquired by the students.

The structure defined above outlines the reading domain in such a manner that the goals of a given reading program may be explicated in terms of the behavioral objectives of the PROBE system. In order to measure the attainment of these goals, item pools will be developed for each behavioral objective. Assessment devices may then be generated by selecting test items from the pools corresponding to the objectives of the program. In this manner, a set of parallel tests may be generated for any combination of objectives desired by the user.

In addition to the reading structure and the test items, the system will include a user's guide, which will explain procedures whereby a user may use SOBE-R to define his program goals in terms of behavioral objectives, and ultimately to obtain assessment devices to be used in the evaluation of the program.

Completed. Methodologically, SOBE-R may be the prototype for an entirely new method of developing assessment devices, i.e., tests. The structure of the reading domain has been defined in such a way that the method should be directly applicable to other content areas. Several advantages may be gained as a result of expressing a content area in terms of the structure.

First, the level of specificity attained makes it possible for close scrutiny to determine whether certain important areas have been omitted from the structure. This allows for improvement of the system by filling in those gaps created by the omission of important objectives. Present techniques for categorizing content area for use in test development use such large categories that one cannot be sure that the domain of the content area is comprehensively covered by the test items.

Second, the assessment of learning may be more accurately accomplished using tests derived from such a structure. The flexibility of the system allows the user to produce a test which is tailor-made to fit the particular objectives of the program.

Third, the scores of a test produced from such a system are more meaningful. Since the items are related to behavioral objectives, the test is truly criterion-referenced. That is, the score on the test is an indicator of the extent to which the learner has achieved the objectives, rather than an ordinal position with respect to the scores of a group of his peers. Depending on the extent to which such knowledge is desired, the test score may be broken down into sub-scores which may be indicative of very specific types of behavior. This procedure may be particularly useful in diagnostic testing.

Several products either have or will come out of the PROBE project. A major product is the reading structure. A first draft of the structure has been completed, and will be reviewed by reading experts with respect to its accuracy, completeness, etc. This structure comprehensively defines the reading domain in terms of behavioral objectives beginning with pre-reading skills through reading skills in content areas, which would be applicable in the secondary schools.

Planned. Another major product which will be developed is a set of item pools for the behavioral objectives. Several items will be written for each objective to constitute the item pool for that objective. Work on this has already begun.

Other products are also necessary to make this a complete system. For example, user's guides will be produced which will permit the implementation of the system outside CSE. These user's guides will describe such things as how to locate the objectives which define the user's program, how to synthesize objectives of different interest groups into a single, unified, i.e., consensus, set of objectives. Also included in the guides will be procedures for developing summary tests for those objectives selected.

Another product which will emerge from this activity is an index of currently available standardized reading tests with respect to the objectives in the SOBE-R structure. This indexing procedure is currently in progress.

Project Professional Staff

Eva Baker
Ted Dahl
Julie Fitzgerald
J. Ward Keesling
Catherine Masi
Marianne Patalino
W. James Popham
Estelle Shane
Rodney W. Skager, Project Director
Gretchen Stangel

* * * * *

PROGRAM ON EVALUATION METHODOLOGY AND THEORY

TRAINING MATERIALS DEVELOPMENT PROJECT

This project is developing and field testing two sets of evaluation training materials. The first set, the *Simulated Evaluation Exercise*, is designed for individuals who are sophisticated in measurement techniques and who desire training in the evaluation process itself. The major goals of this two day exercise are to upgrade the evaluation skills of educational researchers in gathering appropriate information for decision-makers, and to increase the flexibility of evaluators in responding to constraints likely to occur in actual field conditions. To achieve these goals, 4 to 6 man teams of participants play the role of the "evaluator" in an urban remedial mathematics project. The teams are faced with a series of realistic problems regarding the evaluation of this project and they must develop solutions to them. Teams are allowed to branch in the sense that they request and receive information from the file and the cast of characters in the simulated setting. Thus, teams may be operating on different information bases for each problem depending upon their requests. After each problem, however, all teams are given feedback on how the problem might have been solved and brought up to date on the available information. Audio-visual aids and related materials based on actual evaluation studies are used to maintain the realism of the simulation. The exercise does require a leader who is familiar with evaluation, but a detailed "Director's Guide" provides him with the step-by-step procedures he should follow and the points he should cover during the various discussion periods.

The second set of training materials, *Evaluation Workshop I: Orientation*, is designed for school administrators and project directors. Its major purposes are to orient these individuals to the kinds of information an evaluation can and should provide for educational decision-making and to the general procedures and problems involved in selecting, collecting, analyzing, and reporting that information. The specific goals of this exercise are as follows:

1. Describe the major types of evaluation activities and the sequence in which they should be conducted.
2. Know the kinds of information these activities provide as indicated by knowledge of the kinds of data collected; and how that data should be analyzed, reported, and used.
3. Differentiate the responsibilities of the evaluator from those of the project director in each of the various kinds of evaluation activities.
4. Understand the reasons for using various kinds of evaluation techniques, requirements, and methodologies that might be employed in an evaluation (such as tests, control groups, and clearly stated objectives) and recognize when they are used properly versus improperly.
5. Know the factors to consider in planning a project and its evaulation so as to minimize possible problems as indicated by the ability to recognize communications, logistical, and other difficulties in a project plan.
6. Believe that the information provided by a thorough evaluation is worth the time and cost it takes to conduct.

To achieve these goals, the workshop participants form teams of three participants each that play the role of the evaluator in a simulated evaluation of a 10th grade biology course. They are involved, therefore, in conducting the needs assessment, planning the program and its evaluation, determining whether the program was implemented properly, assessing what might be done to improve it and, finally, measuring and reporting upon its effectiveness. The basic instructional model of the workshop consists of presenting instructional materials (via pamphlets, lectures, and audio tapes of phone calls), having the participants do some of the evaluation tasks, and then providing feedback on how these tasks should have been performed. These three phases of *instruction, practice,* and *feedback* are repeated for each of the evaluation activities described above. Since each participant collects his own set of key instructional materials during the workshop, by the end of the session he has a complete, outlined guide to the basic evaluation processes and activities.

The workshop can be conducted for 6 to 200 people, although one coordinator is needed for each group of 12 to 15 participants (4-5 teams). It takes approximately seven hours to run (not including lunch or optional discussion sessions). Thus, it is generally advisable to schedule 1½ to 2 days for it. A Leader's Manual also has been developed which describes the step-by-step details of how the workshop should be conducted including pre-workshop preparations, what should be read to the participants, what tapes to play at what times, etc.

These two sets of training materials represent the Center's initial efforts in developing instructional guides in evaluation. They are "exportable" products in the sense that it is not necessary that they be conducted by Center staff. Further, both sets follow the Center's model and provide a forum in which this model's appropriateness can be tested, discussed, and thereby revised. The training sessions themselves, because of the mix of people they attract and the team nature of the exercises, also provide a good opportunity for evaluators, administrators, and project directors to interact and see each other's problems and viewpoints.

1. *Simulated Evaluation Exercise*

Completed. The exercise and its Director's Guide have been completed. These have undergone two developmental field tryouts and revisions. These field tests were aimed primarily at modifying the initial exercise developed in 1968 to coincide more closely with the Center's model of evaluation and to place greater emphasis on the participants' flexibility in solving evaluation problems.

Planned. The *Simulated Evaluation Exercise* will undergo final field testing this year. The focus of the field tests will be upon double checking its ability to develop flexibility in solving evaluation problems. Measures of the exercise's ability to accomplish this have already been developed and negotiations are now underway to set up the final field tests.

2. *Evaluation Workshop I: Orientation*

Completed. The evaluation workshop has undergone its first set of extensive field tests. These field tests have included two pre-sessions at the California Educational Research Association meeting, as well as tryouts for the San Diego County Supplementary Education Center, the

Northwest Regional Educational Laboratory meeting of research staff and state coordinators, the Orange County Supplementary Education Center, the Plumas Unified School District, the Administrative Cabinet of the Seattle Public Schools, and the Alaska State Department of Education. These field tests were concerned primarily with the appropriateness of the content of the workshop materials. Thus, most of the revisions in the workshop as a result of these field tests dealt with changing the sequence of instructional units, improving the realism of the simulation, and modifying the information, instructions, and forms presented to the participants to improve their comprehension of and interest in the workshop's content. Numerous modifications were made in the workshop after each field test, however, the nature of these changed considerably during the past year (i.e., from complete rewrites and substitutions of whole units to minor editorial revisions within units). In addition to the content of the workshop, the Leader's Manual and the instruments used to evaluate the exercise also were field tested and revised.

The content of the workshop is now in relatively final form and the achievement test results from its most recent field tests have indicated that it is meeting its objectives when it is conducted properly. Questionnaire data also indicate that the participants felt that the workshop was well run and valuable to them in their jobs as project directors and school administrators. To insure final implementation of the workshop materials, the Center has arrived at an agreement with the Northwest Regional Educational Laboratory (NWREL). The terms of this agreement include the Center's continued development of evaluation training materials and the refinement of them on the basis of field tests that NWREL will continue to set up and coordinate.

Planned. The final field tests of *Evaluation Workshop I* will focus upon a streamlined format to facilitate publication and distribution as well as to reduce administration time. Discussions with potential publishers are already underway regarding the nature of this format. Further, several of the final field tests have already been scheduled to begin August of 1970 and to be completed by January 1, 1971.

Project Professional Staff

James Burry
David Churchman
Stephen P. Klein, Project Director
Marc Nadeau

* * * * *

OPPORTUNITY PROJECTS

The Center devotes a small portion of its resources to the development of opportunity projects. These resources provide funding for individually proposed and conducted projects which are likely to contribute to the improvement of evaluation methodology.

1. *Development of Accountability Measurement Procedures*

Completed. Under the auspices of this project, the Center has been working with the Bilingual Programs Branch of the U.S. Office of Education in developing the Independent Educational Accomplishment Audit (IEAA). This new concept in accountability is being implemented this year in Title VII Projects. The Center provided the services of one of its staff members who helped write the guidelines presently being used by teams auditing bilingual projects in California. The Center staff member, Mr. Dale Woolley, is also serving on two audit teams visiting seven different projects in California.

2. *Services to Government Agencies*

Another area of accomplishment has been in providing the services of Center staff to USOE review teams and in consulting services to various government agencies. Several members of the CSE staff have been active in providing consulting services to government agencies and organizations. A partial listing follows:

Completed. Marvin C. Alkin, CSE Director, was a member of the USOE Review Team for assessing the evaluation of the Follow Through Program. The Follow Through Program was established under the Economic Opportunity Act of 1964 and provides for 161 Follow Through projects in 140 communities nationally. In his report to the USOE, Dr. Alkin provided a review and assessment of the evaluation of the program, for which four million dollars have already been expended, and made specific recommendations for change which should be enacted before additional funds are committed to further evaluation.

Dr. Stephen Klein has served as a consultant on several occasions for the Bureau of Educational Personnel Development of the Office of Education. His consultation has included reports and on-site visits dealing with how the Bureau might evaluate its major programs and the projects in these programs. He has also devoted some of his time and evaluation expertise to the Direction Sports project, an after school tutorial program in reading and mathematics for disadvantaged Black and Brown urban youth. This project has developed to the point where it is now being included in the Model Cities Program.

Dr. C. Robert Pace has been active as a reviewer of research proposals for the U.S. Office of Education and for the National Research Council. He also worked as a consultant for a background paper on higher education for the White House National Goals Research Staff. Dr. Pace was invited by the White House to submit a paper entitled "Adaptation of Higher Education to Individual and Societal Needs and Historical Perspectives."

Dr. James W. Trent has been involved in extensive consulting activities during the past year. He has been a consultant for the Bureau of Higher Education, involved in planning a 3 million dollar longitudinal study, and for the American Association of Junior Colleges to plan needs in research in higher education and junior colleges.

Dr. W. James Popham consulted with the California Joint Assembly-Senate Committee on Educational Goals and Evaluation.

3. *Matrix Sampling*

One further concern of the Center has been with the assessment of multiple outcomes of educational programs and the development of techniques for measuring these outcomes. The simple provision of tests, however, does not solve the problem adequately if there are no practical methods for obtaining scores. The Center believes that the technique of matrix sampling may be one effective answer to this problem.

In matrix sampling the basic idea is that it is not necessary for every student taking an examination to respond to every item in the test in order to obtain estimates of

the mean and variance of the examinees' response to the items. In terms of the evaluation of instructional programs, the evaluator wants to know how a group of students performs on a group of items; the technique of matrix sampling provides the evaluator with an effective means of obtaining that information.

Matrix sampling involves the simultaneous, random sampling of both students and items. One matrix sample is a sample of students taking a sample of items. The most efficient use of the technique involves different, non-overlapping samples of students taking non-overlapping samples of items. That is, one sample of students takes one sample of items, and so forth.

Completed. The work of the Center in matrix sampling in the past year has been conducted at both the theoretical and empirical level. CSE Report No. 5 places matrix sampling in context, as one of several kinds of inference-making procedures. The paper also discusses basic formulae and the role of matrix sampling in educational research. CSE Report No. 52 is concerned with testing the technique empirically by contrasting it with examinee sampling. One finding was that the theory could be conceptually simplified by casting the technique into an analysis of variance framework. *Evaluation Comment* 1, 3 provides several definitive statements on the uses and cautions of the matrix sampling technique.

The technique of matrix sampling is of potential value for both the typical classroom situation and the large scale evaluation program. In both situations the technique allows the exploration of a number of new and potentially important variables related to the outcomes of educational programs.

Although the nature of the Opportunity Project is such that there are no specific staff assignments, the following members of CSE have been actively involved in the project:

Marvin C. Alkin
Gary Fenstermacher
Stephen Klein
C. Robert Place
W. James Popham
Allan Rosenstein
Richard Seligman, Project Director
Kenneth Sirotnik
Rodney W. Skager
James W. Trent
Dale Woolley

* * * * *

IMPACT ON NEW PROGRAMS AND SERVICES

There are a number of activities that have been engaged in by Center personnel which constitute an outgrowth of CSE activities. While these efforts are not supported by the Title IV grant constituting the Center for the Study of Evaluation, they nonetheless represent significant achievements and contributions of the Center. Many of these new programs and services have already been mentioned within the context of this report and thus will simply be listed here. They are:

1. The Instructional Objectives Exchange, Inc. — A national clearinghouse for objectives and items providing services to teachers, school districts, state departments, etc.

2. Education for the Handicapped: Evaluation Project — A national R & D Center, supported by the Bureau of Education for the Handicapped, Office of Education, has been established to provide working models, systems, and procedures for the evaluation of handicapped programs and students at the preschool level.

3. Planning grant for the development of a proposal for training product developers and evaluators — This activity ties into both the PROBE project of the Center and the Evaluation Training Materials project. Four of its five principal staff members are CSE personnel.

4. Evaluation materials for the ES-70's program — This project is now completed but was run in conjunction with ongoing Center activities.

5. A Study of Junior Colleges — Another study conducted by CSE staff under separate USOE funding; it deals with an evaluation of the impact of 2-year institutions on the social, economic, educational, and manpower needs of the communities served, and other related matters.

6. An Analytical Review of Longitudinal and Related Higher Education Studies as They Apply to the Educational Process — A higher education evaluation study being directed by CSE staff funded by USOE.

7. Close working relationships with commercial test publishers and impact on improving the quality of their materials.

ADMINISTRATION

The Center is now in its third year under the directorship of Dr. Marvin C. Alkin, who is also a faculty member of the Graduate School of Education. A Management Advisory Committee composed of elected senior research staff members provides the Director with a highly coordinated team which both recommends and implements program policies. The Assistant Director of the Center, Dr. Richard Seligman, is responsible for supervising central service functions of the Center, for monitoring several Center activities, and for maintaining the various liaison relationships.

The Center is not unwilling to follow its own preachings with respect to the nature of the evaluation function. Fortunately, the U.S. Office of Education has served the Center well by providing regular site review teams to comment on the over-all progress of CSE. While this group has provided some implementation and progress evaluation information, it must be viewed primarily as oriented toward providing an outcome evaluation.

In order to provide for regular and continuing evaluation information to insure program modification as needed, CSE has appointed Dr. Garth Sorenson, a faculty member of the UCLA Graduate School of Education, as the program evaluator for the Center. He is not engaged in any of the CSE projects and is charged with the responsibility of providing evaluation information to the Center Director.

SUMMARY

The National R & D Center Program has demonstrated its worth many times over. Investment in the R & D Center program has yielded usable products which are field relevant and which contribute to the improvement of education. Based on these criteria, the Center for the Study of Evaluation, in terms of its significant contributions to improving the efficiency of American education, clearly rates as one of the leaders. CSE has been fulfilling its goal of producing products, procedures, and methodology which are directly usable for improving the educational accountability of schools. The Center welcomes the continuing challenge to do even more in this area.

Statement of Intent

The Center for the Study of Evaluation, founded in June, 1966, is an educational research and development center sponsored by the U.S. Office of Education under Title IV of the Elementary and Secondary Education Act of 1965.

The Center, directed by Marvin C. Alkin, is a unique organization working exclusively on problems of educational evaluation and is devoted to three prime objectives: to develop a theory for the study of evaluation; to develop methods and instruments for measuring program effectiveness; and to provide a scientific basis for program and policy decisions in education. After an initial period of exploration, the Center's efforts have been increasingly focused on a relatively few research projects which fall within the scope of major program areas. This programmatic approach has resulted in combined efforts by specialists in various disciplines and in the development of a conceptual framework around which a comprehensive theory of evaluation can be built.

Evaluation Comment provides a forum for the discussion of significant ideas and issues in the study of evaluation of educational programs and systems. *Evaluation Comment* is especially interested in publishing creative or controversial ideas, concepts, and dialogue about evaluation of instructional programs that promise to improve knowledge about evaluation or, at least, to excite interest and comment from readers.

A copy of *Evaluation Comment* is distributed free of charge to each scholar, researcher, or practitioner on our mailing list. One to five additional copies may be obtained free of charge; however, where greater amounts are needed readers are encouraged to reproduce the *Comment* for their own purposes. To be included in our mailing list or to order, subject to availability, additional copies of *Evaluation Comment*, please write to:

James Burry, Managing Editor
Evaluation Comment
Center for the Study of Evaluation
145 Moore Hall
University of California, Los Angeles
Los Angeles, California 900[illegible]

Marvin C. Alkin, Editor

UCLA CSE Evaluation comment
Center for the Study of Evaluation

Center for the Study of Evaluation
145 Moore Hall
University of California
405 Hilgard Avenue
Los Angeles, California 90024

APPENDICES

APPENDIX A

Summary of Center Accomplishments

SUMMARY OF CENTER ACCOMPLISHMENTS

A. School Evaluation Project
   1. Elementary School Evaluation KIT
   2. Preschool-Kindergarten School Evaluation KIT
   3. Attitude-Toward-School Questionnaire
   4. CSE Elementary School Test Evaluations
   5. Science Test Prototype
   6. Atlas of Scores

B. Higher Education Evaluation Project
   1. Higher Education Questionnaire
   2. Measures for Higher Education Evaluation Kit

C. PROBE
   1. Instructional Objectives Exchange (IOX)
   2. Objectives Collections
   3. National Network of Schools
   4. System for Objective-Based Evaluation - Reading (SOBE-R)

D. Training Materials Development Project
   1. Simulated Evaluation Exercise
   2. Evaluation Workshop

E. Opportunity Projects
   1. Independent Educational Accomplishment Audit
   2. Consulting Services to Government Agencies
   3. Matrix Sampling

APPENDIX B

CSE Research Staff

CSE RESEARCH STAFF

MARVIN C. ALKIN, CSE Director, helped write the original research proposal as a member of the Center's founding committee. He is an associate professor of education at UCLA. Professor Alkin received his doctorate from the Stanford University School of Education. His papers include such subjects as mathematical models for school district evaluation, data accessibility in school district research, economy of scale in education, and cost effectiveness evaluation of education. His most recent work is related to the development of evaluation theory.

EVA L. BAKER is a member of the PROBE staff. An assistant professor at UCLA, she specializes in teacher education, instruction, and product research. She received her Ed.D. from UCLA and was employed as a member of the professional staff of the Southwest Regional Laboratory for Educational Research and Development. Her major research interests involve determining the contingencies under which program evaluation results in program improvement. Her responsibilities on the project center on implementation of the system. She is also associated with the Instructional Objectives Exchange, an agency funded originally by the Center.

PAUL BRADLEY is an assistant research educationist for the School Evaluation Project. He received an A.B. from Brown University and a Ph.D. in psychology from the University of Southern California. Dr. Bradley was a research assistant for the Aptitudes Research Project, where he was engaged in test construction and data analysis. Prior to joining the Center he was an assistant professor of educational psychology at New York University.

JAMES BURRY is Managing Editor of the Center's Evaluation Comment and a member of the staff of the Training Materials Development Project at the Center. He holds a B.A. and M.A. in English and is presently a candidate for the Ph.D. at UCLA, specializing in literature and linguistics. In the past he has written for newspapers in Europe and the United States, and has worked as an editor in the fields of education and English, and as translator in several languages.

DAVID CHURCHMAN, a member of the Training Materials Development Project, received his B.A. and M.A. from the University of Michigan, specializing in Near Eastern history, and is now a candidate for an Ed.D. in educational evaluation at UCLA. He has experience as a librarian in Michigan, a social worker in California, and has taught in California, Pennsylvania, and Morocco. Mr. Churchman's publications include an earlier exercise for training evaluators and articles on travel, personnel selection and curriculum. Currently his major project is an evaluation of American-sponsored schools in Europe and the Near East, under the auspices of the American Association of School Administrators.

THEODORE A. DAHL, Coordinator of Research for the Project for Research of Objective-Based Evaluation, is a doctoral student in the UCLA Graduate School of Education. He received a B.S. in engineering from UCLA, where he participated in the evaluation of prosthetic devices. His Ph.D. area is measurement and statistics. Mr. Dahl was a member of the staff at the Southwest Regional Laboratory for Educational Research and Development.

WILLIAM DOHERTY serves the School Evaluation Project on the preschool and kindergarten evaluation phase. He is presently in the doctoral program at the University of Southern California in quantitative psychology. Mr. Doherty's previous experience includes working at the Aptitudes Research Project where he was research assistant on the Figural Evaluation Study.

JAMES S. DYER, an assistant professor in the Graduate School of Business Administration, is interested in the application of systems analysis and operations research approaches to the problems of education. He teaches courses at UCLA in the areas of mathematical programming, project management, and public systems analysis. He also serves as a consultant to the RAND Corporation on problems relating to the use of PPBS and other methods of analysis in education. Dr. Dyer's dissertation was entitled Cost-Effectiveness Analysis for a Public System of Higher Education.

MARSHA EHRENBERG is a member of the Higher Education Project. She obtained an M.A. degree in higher education from UCLA. Her special interest is innovation in higher education. Miss Ehrenberg's primary duties on the project include background research and data analysis.

GARY FENSTERMACHER currently serves the Center as an advisor on evaluation theory. He completed his Ph.D. in Philosophy of Education at Cornell University, and has taught at New York and Hofstra Universities. Professor Fenstermacher was trained in the school of philosophy known as conceptual analysis, and is interested in the relationships between behavioral science theory and educational practice. In conjunction with his Center colleagues, he is developing an evaluation theory which would serve as a model for evaluation strategies.

JULIE FITZGERALD is a member of the PROBE research staff. She has considerable experience in the area of objective-based teaching and evaluation, gained through her former position on the staff of the Instructional Objectives Exchange. In the 1969 IOX Summer Institute, Mrs. Fitzgerald directed a group of teachers in the writing, editing, and modification of behavioral learning objectives and associated test items.

RALPH HOEPFNER, an associate research educationist, is a measurement and evaluation specialist directing the School Evaluation Project. He holds a Ph.D. in psychological measurement from the University of Southern California and has been engaged in research on measuring differential aspects of intelligence and their relationship to academic success. In his present position he has developed critical evaluation methods for educational tests and systems for internal evaluation of schools and districts.

MARVIN HOFFENBERG is a member of the School Evaluation Project. Currently he is also a lecturer in political science at UCLA and a consultant for the Institute for the Future and for RAND Corporation. Previously, he has served as a research economist, research director for several federal agencies, both in the United States and abroad, and for 11 years was the Assistant Chief of the Division of Interindustry Economics in the U.S. Bureau of Labor Statistics. His area of major interest is economics and he specializes in input-output theory and application and in cost analysis. He has published extensively in these areas throughout the last 25 years.

SONJA JACOBSON, Project Manager of the Higher Education Project, is currently working toward her doctorate in higher education at UCLA. She received her B.A. in social psychology from the University of California, Berkeley, and did graduate work in educational psychology at the University of Southern California.

PATRICE JANSEN is a research assistant for the School Evaluation Project. Her major concern on the project is the evaluation of tests. Miss Jansen obtained her master's degree from UCLA in public health with an emphasis in statistics. She has previous experience in technical writing for the aerospace industry.

J. WARD KEESLING, a member of the PROBE staff, is an assistant professor of educational measurement and research design. He is completing his dissertation in the Measurement Evaluation and Statistics Analysis program of the Department of Education, University of Chicago. While at Chicago, he served as a research assistant in the Early Education Research Center. In 1968, Mr. Keesling served as consultant in statistics and data processing at a seminar on "Learning and the Educational Process" at Stockhlom, Sweden. His research interests focus upon making causal inferences from non-experimental data and measurement models suitable for use with criterion-referenced tests.

STEPHEN P. KLEIN, Project Director for the Evaluation Training Materials Project, received his doctorate in psychology from Purdue University and has had four years experience as a research psychologist with the Educational Testing Service. Dr. Klein has published numerous works in the areas of creativity, educational measurement, and criterion development and has taught measurement courses at the UCLA Graduate School of Education, conducted post-doctoral training sessions in evaluation, and done extensive consulting in evaluation problems for several school districts and educational agencies, e.g., the Department of Health, Education, and Welfare; the Graduate Record Examination Board; and the Educational Testing Service.

DON LONG, Senior Programmer for Statistical Services at the Center, acquired his theoretical background from Columbia University and his technical background from I.B.M. Corporation. Mr. Long also has diversified data processing experience from the physics and sociology departments at UCLA. In the past, he has been a lecturer in statistics for the sociology department and has been a high school math teacher.

ANN I. MOREY, Executive Officer for the Higher Education Evaluation Project, received her doctorate in higher education from the University of California, Berkeley. Her research interests have focused on studies of institutional images, student self-selection, and college persistence. Dr. Morey has previously been affiliated with the Division of Educational Statistics, U.C. Berkeley, as coordinator of the junior college project, with Newark State College as assistant professor of education and administration, and with the Pennsylvania State University as a staff member of the Dean of Women's Office.

MARC-ANDRE NADEAU is working on the development of training materials for the simulated evaluation workshops. He received his B.A. and M.A. from Laval University in Quebec, specializing in measurement and testing, and is currently a candidate for the Ph.D. in the Graduate School of Education, UCLA. Mr. Nadeau is an assistant professor in measurement and statistics at Laval University (on leave). His publications include articles on learning to read and reading readiness. His special interests are evaluation and training.

C. ROBERT PACE, director of CSE projects on the development of contextual and criterion measures for higher education, has devoted his major research interests in higher education to evaluation and measurement, college environments and follow-up studies of college graduates. He received his M.A. and Ph.D. in educational psychology from the University of Minnesota. A professor of higher education at UCLA since 1961, Dr. Pace's previous affiliations were with the Evaluation Center and the Psychology Department at Syracuse University, the Personnel Research Section of the Bureau of Naval Personnel, and the American Council on Education.

MARIANNE PATALINO is currently working on the Project for Research on Objective-Based Evaluation, developing a structure for the classification of reading skills and objectives. She received her B.A. in French and M.A. in measurement; both degrees are from UCLA. She has been with the Center for the Study of Evaluation since 1967.

W. JAMES POPHAM, professor at UCLA and member of the PROBE staff, is deeply involved in various aspects of educational technology. His most recent research has been in the evaluation of programmed instructional materials for teacher education. In the past five years he has also conducted U.S.O.E.-sponsored projects to develop performance tests of instructor competence. He taught at San Francisco State College and at Indiana University, where he earned his Ed.D. in secondary education. Dr. Popham is also associated with the Instructional Objectives Exchange.

ALLEN B. ROSENSTEIN serves the Center as a staff member involved in the conceptualization of evaluation. He also is a professor in the School of Engineering. Dr. Rosenstein has served as a consultant to more than 30 corporations and government agencies during the last 25 years. He has wide experience in teaching and has served as a faculty lecturer and consultant regarding education and curriculum design for 35 universities and technical institutes both in the United States and abroad. For the past 12 years, until its recent completion, Dr. Rosenstein served as the Principle Investigator for an extensive study of professionals and education for the professions.

RICHARD SELIGMAN, Assistant Director of CSE, is a specialist in the social-psychological study of higher education. Associated with the Center since 1966, Dr. Seligman was previously involved in the Higher Education Project. He holds an M.A. from Ohio University and an Ed.D. from UCLA's Graduate School of Education. His research interests have focused on studies of the college environment--specifically on the development of contextual measures of various attributes of institutions of higher learning.

ESTELLE SHANE is associated with PROBE. She has been working on the problems of implementing an objective-based evaluation system and has helped coordinate the Network of Schools activities. Mrs. Shane has been advanced to candidacy for her doctorate in curriculum and instruction. She received an M.A. in English from UCLA, has served as a visiting professor at California State College at Long Beach, and is currently teaching English at Valley Junior College.

RODNEY W. SKAGER, Project Director of PROBE, has been with CSE since its inception in 1966. His recent work has centered on the development of systems for utilizing sets of behavioral objectives and associated test item pools in the evaluation of instruction. He holds an M.A. and Ph.D. in psychology from UCLA. An associate professor of education and a research psychologist, his research interests have centered on the evaluation of instructional programs, the measurement of cognitive development, and the function of educational research in the process of making decisions in education. The latter area is the subject of a book now in press.

GARTH SORENSON, Professor of Educational Psychology and Counseling Theory at UCLA's Graduate School of Education, serves as program evaluator for the Center. Professor Sorenson does not participate directly in CSE projects but, rather, holds the responsibility of providing evaluation information on CSE activities to the Director. In his 15 years at UCLA, Dr. Sorenson, who holds advanced degrees in sociology and educational psychology from the University of Utah, has conducted research in teacher education, counselor training, and counselor effectiveness. Dr. Sorenson serves as a consultant to various public and private schools in the Los Angeles and San Diego areas. In recent years, he has been assisting several Regional Educational Laboratories in evaluating their activities.

GRETCHEN STANGEL is a member of the PROBE research staff. She received her M.S. degree in educational psychology from the University of Wisconsin. While on the staff of the School Evaluation Project, Mrs. Stangel conducted an extensive review of standardized tests for the Elementary School Test Evaluations. Her knowledge of the content, strengths and weaknesses of standardized tests is augmented by considerable experience in test item writing and in test administration.

GUY STRICKLAND, a member of the staff of the School Evaluation Project, has been concerned with test evaluations and with cost-effectiveness. He received his B.A. in economics from Williams College. His M.A. thesis in education dealt with attitudes toward school in the primary grades. Mr. Strickland has five years experience as an elementary schoolteacher. He is currently a student in the UCLA Graduate School of Business and is working toward a doctorate in educational evaluation.

JAMES W. TRENT, an associate professor at UCLA, serves the CSE in projects relating to higher education. He is working to develop criterion and contextual measures for this field. Previously affiliated with the U.S.O.E. R & D Center at Berkeley, Dr. Trent has written extensively on psychosociological studies of high school graduates, including studies of characteristics associated with various patterns of college attendance, attitude changes of college and non-college peer groups, the intellectual development of college students of different religious subcultures and studies of activism and apathy among college students. Dr. Trent has served as a consultant on student development, graduate education planning, research in multiple-media teaching techniques, self study projects, junior college programs for cultural minorities, and U.S.O.E. program planning.

KENNETH A. TYE is serving CSE as an advisor to the School Evaluation Project. Currently, he is also the Assistant Director of the Institute for the Development of Educational Activities in Los Angeles. Dr. Tye received his M.A. in elementary administration from San Diego State College and his Ed.D. from the School of Education at UCLA. Dr. Tye has toured the educational systems of Europe for the U.S. Office of Education and the Ford Foundation and also participated in the Teacher Training Program of "Operation Crossroads Africa" in Liberia, West Africa, in 1962. In addition to studying in the area of African educational problems, Dr. Tye has published a large body of material on the problems of elementary school administration.

LOUISE TYLER, an associate professor of curriculum and instruction in UCLA's Graduate School of Education, is a member of the research staff on the School Evaluation Project. She received an M.A. in educational psychology and Ph.D. in curriculum and evaluation from the University of Chicago. Dr. Tyler is a consultant for the research division of the Institute for the Development of Educational Activities. Her central interests lie in the field of curriculum construction and evaluation, drawing upon the theoretical framework of psychology and related disciplines for investigations in the field.

DALE WOOLLEY, Project Manager for the School Evaluation Project, is presently responsible for field-test coordination with cooperating school districts and relations with test publishers. His previous research activities include the development of staff performance evaluation systems, school-community communication models, and procedures for implementing the concept of educational auditing. Before joining the Center staff last year, Mr. Woolley had completed ten years as teacher, counselor and administrator in the public schools. He has completed degrees at the University of California at Davis and the University of Southern California.

APPENDIX C

PUBLICATION LIST

PUBLICATION LIST

CSE REPORTS ERIC Numbers

CSE Report No. 1; 22pp; Nov., 1966
CATEGORIES OF COGNITIVE SKILLS ED 013 974
Harriet Foster
This paper is an attempt to identify the significant aspects of cognition as one step in the measurement of cognitive skills. It is concerned with identifying dimensions of cognitive tasks which may be relevant to questions of generalizing research findings, and to measuring generalizations of learning and performance.

CSE Report No. 2; 89pp; Dec., 1968
MANUAL OF INDIVIDUAL DIFFERENCE VARIABLES AND MEASURES ED 036 877
Norma D. Feshbach
The manual offers a source of variables of potential use as behavioral criteria in an evaluation program. Twenty-nine individual difference variables are included and briefly defined. The measures used to assess the variables are described, and evaluation of the variables and pertinent references are listed. Each variable and its definition is also available separately.

CSE Report No. 3; 9pp; Dec., 1966
THE EXPERIMENT IN RESEARCH ON EVALUATION OF INSTRUCTION ED 012 107
Merlin C. Wittrock
The study reported here was concerned with the evaluation of contingencies between instructional variables and their multiple effects. The objective was to provide an instrument for securing knowledge about the process of evaluation, and to define those areas of instruction wherein evaluation is most effective.

CSE Report No. 4; 21pp; Feb., 1967
CLOZE READABILITY PROCEDURE ED 010 983
John R. Bormuth
This report examines the Cloze procedure for evaluating the comprehension difficulty of written instructional materials. Research bearing on the validity, the formal characteristics, and the application of the procedure are reviewed. A bibliography is included.

CSE Report No. 5; 31pp; Dec., 1967
ITEM SAMPLING IN EDUCATIONAL RESEARCH ED 013 975
T.R. Husek and K. Sirotnik
In its discussion of item sampling and its uses, this paper considers the possible cases that can arise when making inferences from a sample of examinees and/or test items to a population of examinees and test items. Formulas are provided so that the data from the item sample can be used to estimate the mean and variance of the randomly sampled population.

CSE Report No. 6; 29pp; Feb., 1967
INPUT-OUTPUT RELATIONSHIPS IN A SAMPLE OF CALIFORNIA PUBLIC JUNIOR COLLEGES ED 014 989
Marvin C. Alkin & Vernon L. Hendrix
With statistical control for variations in community characteristics, 15 public California junior colleges were studied with respect to several financial variables and output measures. In this report, variation in the output measures and in the financial input are explained in detail. The effects of community differences are discussed.

CSE Report No. 7; 53pp; Dec., 1968
PROGRAM BUDGETING
John Hagen
This report synthesizes current knowledge of program budgeting and provides comprehensive information on budgeting practice and a review of its history and relevant literature. The study provides familiarity with the various concepts involved, along with a guide to available data.

CSE Report No. 8; 20pp; Dec., 1968
EVALUATION PERSPECTIVES: 1968
C. Robert Pace
This paper discusses the historical lines of development of evaluation and describes the role of the evaluator in terms of that development. The study is concerned with emerging emphases in evaluation and the applicability of different evaluation models to different units which may be evaluated.

CSE Report No. 9; 41pp; Aug., 1968
(From the CSE Symposium on Evaluation)
TOWARD A THEORY OF TESTING WHICH INCLUDES MEASUREMENT-EVALUATION-ASSESSMENT
ED 036 878
Benjamin S. Bloom
The main thesis of this paper is that testing is now ready for a major effort to create a synthesis out of what has hitherto been a series of unrelated approaches to testing. The paper suggests ways in which some of the powerful aspects of each different testing approach--measurement, evaluation and assessment--may be brought together into a more complex and useful way of handling test problems.

CSE Report No. 10; 9pp; Sept., 1968
(From the CSE Symposium on Evaluation)
EVALUATION AS A MAIN AIM OF SCIENCE: COMMENTS ON PROFESSOR BLOOM'S PAPER ENTITLED 'TOWARD A THEORY OF TESTING WHICH INCLUDES MEASUREMENT-EVALUATION-ASSESSMENT'
Michael Scriven
In this discussion, Scriven suggests a broader perspective of evaluation than that offered by Bloom and proposes a more conservative and less technical taxonomy for the scientific investigation of education.

CSE Report No. 11; 11pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR BLOOM'S PAPER ENTITLED 'TOWARD A THEORY OF TESTING WHICH INCLUDES MEASUREMENT-EVALUATION-ASSESSMENT'
Gene V. Glass
The question of whether measurement, evaluation, and assessment should be synthesized into a theory of testing is answered negatively by Glass, who feels that such a synthesis would misdirect the development of one of the constituents - evaluation. The paper also offers a conception of evaluation in and of itself.

CSE Report No. 12; 8pp; June 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR BLOOM'S PAPER ENTITLED 'TOWARD A THEORY OF TESTING WHICH INCLUDES MEASUREMENT-EVALUATION-ASSESSMENT'
ED 036 871
J.P. Guilford
This commentary analyzes Bloom's definitions of kinds of evaluation and the needs for evaluation in education. In a discussion of the nature

of tests of cognition, memory, and production and evaluation abilities, Guilford stresses the need for acquisition of specific information and for general intellectual skills to deal with that information.

CSE Report No. 13; 27pp; Sept., 1968
(From the CSE Symposium on Evaluation)
EVALUATION OF INSTRUCTION AND CHANGING EDUCATIONAL MODELS ED 032 647
Robert Glaser
This paper is concerned with the nature and organization of educational practice. Glaser discusses certain trends in current educational practice and suggests a general model for an adaptive instructional system, giving particular attention to probable influences on the evaluation of instruction.

CSE Report No. 14; 8pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR GLASER'S PAPER ENTITLED 'EVALUATION OF INSTRUCTION AND CHANGING EDUCATIONAL MODELS' ED 036 872
Robert Stake
Stake asserts in this paper that behavioral specification, as it is usually defined, is often impractical and unnecessary, and he challenges the primacy of goal specification as a first logical step in instruction or evaluation. Objectives must be reported, he feels, but attention must be given to how they change over time and to the priorities given by different groups to different objectives.

CSE Report No. 15; 7pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR GLASER'S PAPER ENTITLED 'EVALUATION OF INSTRUCTION AND CHANGING EDUCATIONAL MODELS' ED 036 867
Arthur A. Lumsdaine
While in general agreement with Glaser on the importance of behavioral objectives, Lumsdaine stresses the need for providing an explicit rationale or foundation for evaluation. Emphasis is placed upon improving the technology of evaluation and creating a better market for assessment data and data standards.

CSE Report No. 16; 34pp; Sept., 1968
(From the CSE Symposium on Evaluation)

INSTRUCTIONAL VARIABLES AND LEARNING OUTCOMES ED 036 866
Robert M. Gagné
This paper describes the operations which underlie the measurement of educational outcomes. Procedures of laboratory psychologists in achieving distinct criteria for measuring such outcomes as conditioned responses, chains, and multiple discriminations are described. One technique described is two-way measurement, in which the simpler outcome is measured first, then the more complex.

CSE Report No. 17; 11pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR GAGNÉ'S PAPER ENTITLED
'INSTRUCTIONAL VARIABLES AND LEARNING OUTCOMES' ED 036 868
Richard Anderson
In commenting on Gagné's remarks on learning outcomes, Anderson offers a research design appropriate for instructional research studies. He challenges certain assumptions concerning correspondences between subject matter and categories, and asserts that the logic of experimental analysis should be applied to instructional research.

CSE Report No. 18; 9pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR GAGNE'S PAPER ENTITLED
'INSTRUCTIONAL VARIABLES AND LEARNING OUTCOMES' ED 036 873
Leo Postman
Postman stresses here that the categorization of outcomes and measuring operations should be regarded as flexible heuristic devices, and that standardization should be guarded against. He also discusses the classes of outcomes distinguished in experimental investigation and the selection of operations.

CSE Report No. 19; 27pp; Sept., 1968
(From the CSE Symposium on Evaluation)
THE CRACKED CAKE OF EDUCATIONAL CUSTOM AND
EMERGING ISSUES IN EVALUATION ED 036 875
Dan Lortie
This paper examines the process of change in American public schools which results in shifts in assignation of value. Lortie discusses new methods of instruction and larger-scale structural changes which have important consequences for evaluation, and suggests a conception of evaluation that is more expeditious in creating solutions to the problems brought about by change.

CSE Report No. 20; 5pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR LORTIE'S PAPER ENTITLED
'THE CRACKED CAKE OF EDUCATIONAL CUSTOM AND
EMERGING ISSUES IN EVALUATION' ED 036 874
C. Wayne Gordon
This paper suggests that the variety of decision making proposed by Lortie will not afford the luxury of evaluative systems of the kind he describes. Gordon feels that, had Lortie pursued a line of functional analysis of many outcomes, he would have arrived at an entirely new analysis of the justification for complexity in evaluation.

CSE Report No. 21; 9pp; Sept., 1968
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR LORTIE'S PAPER ENTITLED
'THE CRACKED CAKE OF EDUCATIONAL CUSTOM AND
EMERGING ISSUES IN EVALUATION' ED 036 876
N. L. Gage
In examining some of Lortie's assumptions, Gage questions that each school district needs independent evaluation and that evaluation should be applied at the end of a given educational program. The paper offers the possibility of a more sensitive evaluation which seeks to understand more immediately realizable effects.

CSE Report No. 22; 36pp; May, 1969
(From the CSE Symposium on Evaluation)
THE CRITERION PROBLEM IN THE EVALUATION OF INSTRUCTION: ASSESSING POSSIBLE, NOT JUST INTENDED OUTCOMES ED 030 987
Samuel Messick
This paper discusses cognitive styles and affective reactions as two major classes of criterion variables that should be taken into account in the evaluation of instruction. These variables are emphasized because of their bearing upon questions that stem from particular views about the diversity of human performance and the role of values in educational research.

CSE Report No. 23; 4pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR MESSICK'S PAPER ENTITLED
'THE CRITERION PROBLEM IN THE EVALUATION OF
INSTRUCTION: ASSESSING POSSIBLE, NOT JUST
INTENDED OUTCOMES' ED 030 980

Paul Blommers
While Blommers objects to little in Messick's treatment of affective variables, he suggests that Messick's argument is better limited to cognitive styles. He also raises several questions that must be answered before decisions regarding the role of cognitive styles in instructional programs can be made.

CSE Report No. 24; 12pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR MESSICK'S PAPER ENTITLED 'THE CRITERION PROBLEM IN THE EVALUATION OF INSTRUCTION: ASSESSING POSSIBLE, NOT JUST INTENDED OUTCOMES' ED 031 802
Leonard Cahen
In addition to discussing the role of value judgments in evaluation, this paper shows the two types of 'unintended' instructional outcomes. Cahen emphasizes the importance of individual difference measures in the cognitive style areas in curriculum research and the need to keep measurement errors at a minimum at the critical positions on individual difference scales.

CSE Report No. 25; 34pp. May, 1969
(From the CSE Symposium on Evaluation)
EVALUATING THE COST-EFFECTIVENESS OF INSTRUCTIONAL PROGRAMS ED 031 818
Marvin C. Alkin
This paper describes how a cost-effectiveness model can be applied in evaluating educational programs and systems, and illustrates the distinctions between cost-benefit analysis and cost-effectiveness evaluation. It outlines the components of a model for cost-effectiveness evaluation in education, and indicates how this model can be used in diverse evaluation situations.

CSE Report No. 26; 11pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR ALKIN'S PAPER ENTITLED 'EVALUATING THE COST-EFFECTIVENESS OF INSTRUCTIONAL PROGRAMS' ED 031 803
Marvin Hoffenberg
Hoffenberg focuses here upon efforts towards rationalization in the decision process, the institutionalization of the search for problem areas, and the beginning of a general theory of organizational behavior.

The paper suggests the need for a general theory of instructional evaluation in order to unify evaluations.

CSE Report No. 27; 5pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR ALKIN'S PAPER ENTITLED 'EVALUATING THE COST-EFFECTIVENESS OF INSTRUCTIONAL PROGRAMS' ED 031 777
John Bormuth
Bormuth agrees that evaluation should play a role in the formation of public policy on education, but expresses doubt that evaluation is sufficiently developed for such a role. The paper stresses the need for a theory of test writing, without which the practical use of evaluation in formation of public policy does not seem possible.

CSE Report No. 28; 18pp; May, 1969
(From the CSE Symposium on Evaluation)
THE DESIGN AND ANALYSIS OF EVALUATION STUDIES: COMMENTS AND SUGGESTIONS ED 030 988
David E. Wiley
In this paper, Wiley makes distinctions among evaluation, assessment, and appraisal, and narrows the definition of evaluation. He outlines the separate elements of evaluation which have been confused in some studies, and relates these elements to design, analysis, and measurement of evaluation.

CSE Report No. 29; 6pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR WILEY'S PAPER ENTITLED 'THE DESIGN AND ANALYSIS OF EVALUATION STUDIES: COMMENTS AND SUGGESTIONS' ED 030 981
Chester Harris
In his concern with identifying the critical issues of evaluation studies, Harris focuses upon the issue of univariate as opposed to multivariate dependent variable studies, the arbitrary selection of instructional packages, the tendency to interpret every study as if it were being conducted for the first time, and the need to give adequate consideration to prior information.

CSE Report No. 30; 4pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR WILEY'S PAPER ENTITLED 'THE

DESIGN AND ANALYSIS OF EVALUATION STUDIES: COMMENTS AND SUGGESTIONS' ED 030 982
Theodore Husek
The major discussion in this paper centers on methods and problems of item sampling. Husek also emphasizes the distribution of scores on tests as well as the mean, more use of the traditional item, reexamination of the items used in evaluation studies, and an examination of new indices.

CSE Report No. 31; 30pp; May, 1969
(From the CSE Symposium on Evaluation)
METHODOLOGICAL PROBLEMS IN THE EVALUATION OF INNOVATION ED 031 820
Martin Trow
The main concern of this paper is with innovations in the curriculum and modes of teaching and learning, rather than with innovations in broader organizational forms. The emphasis is less on the technical problems of evaluative research and more on the characteristics being studied and assessed, as well as on the social context from which they rise.

CSE Report No. 32; 20pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR TROW'S PAPER ENTITLED 'METHODOLOGICAL PROBLEMS IN THE EVALUATION OF INNOVATION' ED 031 819
Eugene Litwak
Litwak proposes two distinct notions of evaluation: a facilitative evaluation which aims at the improvement of everyday operations, and a more formal evaluation of the entire program. Concerning the use of surveys in evaluation, his paper stresses the need to classify events in terms of their complexities and then to designate appropriate methodological procedures.

CSE Report No. 33; 8pp; May, 1969
(From the CSE Symposium on Evaluation)
COMMENTS ON PROFESSOR TROW'S PAPER ENTITLED 'METHODOLOGICAL PROBLEMS IN THE EVALUATION OF INNOVATION' ED 031 778
David Nasatir
This paper suggests that innovations must be considered in light of their impact upon both real and experimental situations. Among the

many factors that must be optimized is the existential quality of life in the classroom, as well as any putative outcomes for which there is no assurance of future realization.

CSE Report No. 34; 109pp; Sept., 1968
APPROACHES TO PROGRAM ACCOUNTING FOR PUBLIC SCHOOLS ED 026 727
Erick L. Lindman, Editor
This report contains brief descriptions of public school expenditure classification systems designed to facilitate program budgeting and cost analysis. Eleven relevant issues are identified in the paper, and arguments pro and con are reviewed.

CSE Report No. 35; 16pp; Feb., 1967
VARIATION IN TEACHERS' REINFORCEMENT STYLE AND IMITATIVE BEHAVIOR OF CHILDREN DIFFERING IN PERSONALITY CHARACTERISTICS AND SOCIAL BACKGROUND ED 012 273
Norma D. Feshbach
The CSE project reported here studied the effects of differing teacher reinforcement behavior on students. The study subjects were middle and lower class male ninth- and tenth-grade remedial reading students. The differences in imitative behavior between the middle class and lower class boys playing the role of teacher are discussed in detail.

CSE Report No. 36; 122pp; Oct., 1968
EDUCATIONAL ACHIEVEMENT AND ASPIRATIONS OF MEXICAN-AMERICAN YOUTH IN A METROPOLITAN CONTEXT ED 028 012
C. Wayne Gordon, Audrey J. Schwartz, Robert Wenkert, & David Nasatir
This report presents the results of an inquiry into the educational problems of Mexican-American youth in Los Angeles. It is primarily concerned with the influence which the characteristics of individual pupils have on the educational effect of the school, and the manner in which the educational system causes outcomes peculiar to Mexican-American pupils.

CSE Report No. 37; 77pp; Feb., 1969
COMPARATIVE VALUES AND ACHIEVEMENT OF MEXICAN-AMERICAN AND ANGLO PUPILS ED 028 873

Audrey J. Schwartz
This paper examines the manner and extent of differences between Mexican-American and Anglo values, and the manner and extent of value differences within the Mexican-American pupil sub-population. It is especially concerned with how these value orientations are related to the academic achievement of Mexican-American pupils.

CSE Report No. 38; 47pp; Apr., 1967
THE IMPLICATIONS AND USE OF CLOZE PROCEDURE IN THE EVALUATION OF INSTRUCTIONAL PROGRAMS ED 012 074
John R. Bormuths
The Cloze Readability Procedure is examined here to determine its utility as a device for evaluating instructional programs. The possibility of developing a method incorporating the Cloze Procedure for making criterion reference tests over verbally presented instruction is also discussed. A summary of Cloze research and a procedure for using the Cloze test and interpreting and reporting Cloze scores are included.

CSE Report No. 39; 19pp; April, 1967
TEACHER APPRAISAL: A MATCHING PROCESS ED 016 299
Garth Sorenson & Cecily Gross
This paper presents a framework for describing the wide variety of data shown to be important to teacher evaluators. Of six postulated categories of expectations people have of teachers, three relate to non-instructional variables and three to instructional variables. The paper describes an instrument to test this model of beliefs about teacher role.

CSE Report No. 40; 68pp; April, 1967
COGNITIVE STRUCTURES AND EDUCATIONAL EVALUATION ED 016 282
Rodney W. Skager & L.A. Broadbent
The use of cognitive measures as criteria in the evaluation of instructional programs is suggested here. A survey of relevant literature to find descriptions of the task and evidence of relationship between such measures and educational, cultural, and other influences is presented.

CSE Report No. 41; 46pp; Feb., 1968
TOWARD AN INSTRUCTIONAL MODEL FOR COUNSELING ED 015 518
Garth Sorenson
Because counseling absorbs much time, money, and personnel, this paper advocates continuous evaluation along the lines applied to instructional

procedures in the classroom. It stresses more precise definition of both dependent and independent variables and the differences among subjects. A partial model of counseling, free from a cognitive point of view, is presented.

CSE Report No. 42; 109pp; Feb., 1969
PRELIMINARY ANALYSIS OF DATA FOR A SECONDARY SCHOOL INPUT-OUTPUT MODEL — ED 030 983
Marvin C. Alkin, Richard Glinski, & Robert Wininger
This is a preliminary report of a project which has as its goal the construction of a mathematical model representing the interrelationships between certain categories of phenomena of the secondary school. It reports on the initial phase of a preliminary analysis of data and provides an overview of the expectations of the final analysis and a summary of the requirements for a full-scale study.

CSE Report No. 43; 21pp; Aug., 1967
TOWARDS AN EVALUATION MODEL: A SYSTEMS APPROACH — ED 014 150
Marvin C. Alkin
Alkin specifies here what should be observed and judged in evaluating instructional programs. A definition of evaluation is proferred, and much of the paper is a discussion and amplification of that definition.

CSE Report No. 44; 149pp; Dec., 1967
PSYCHOPHYSIOLOGICAL CORRELATES OF FEMALE TEACHER BEHAVIOR AND EMOTIONAL STABILITY: A SEVEN-YEAR LONGITUDINAL INVESTIGATION — ED 021 786
William H. Lucio, Marion A. Wenger, & Thomas Cullen
This paper reports on a CSE study of the relationships of individual differences in the personality and functioning of the automatic nervous system to ratings of teacher behavior, emotional stability, and general health.

CSE Report No. 45; 14pp; Mar., 1968
NET-SHIFT ANALYSIS FOR COMPARING DISTRIBUTIONS OF TEST SCORES — ED 023 170
Erick L. Lindman

This paper suggests a technique for analyzing distributions of test scores which involves comparisons of score distributions of different groups of pupils. The aim is to reveal changes in score distribution which may occur when different teaching methods are used.

CSE Report No. 46; 32pp; Apr., 1968
THE MEASUREMENT OF CAMPUS AND STUDENT MORALE ED 021 281
Lora Robinson & Richard Seligman
A morale scale, with items selected from Pace's College and University Environment Scales, is presented here. The paper suggests an approach to the assessment of institutional morale and the conceptual groundwork for its measurement, and identifies a scale of twenty-two items as a means for assessing morale in the college environment.

CSE Report No. 47; 17pp; June, 1968
A THREE-DIMENSIONAL PROGRAM ACCOUNT CLASSIFICATION SYSTEM FOR PUBLIC SCHOOLS ED 023 188
Erick L. Lindman
The three classification dimensions suggested in this report are: (a) type of school, (b) the standard function-object classification, and (c) a scope-of-service classification. The suggested system relates revenues to expenditures, placing emphasis on the net cost of vocational education, compensatory education, special education, and pupil transportation. The system also produces information needed by state legislatures and Congress concerning actual costs of "aided" programs.

CSE Report No. 48; 33pp; Aug., 1968
THREE EXPERIMENTAL MODES OF COUNSELING ED 036 879
Garth Sorenson & Richard K. Hawkins
In this study, an evaluation model was applied to three experimental modes of counseling in an effort to compare their affects on behavior, moods, and feelings about counseling. The goals of the study were to achieve a method of conducting naturalistic counseling interviews while controlling counselor behavior and to construct the necessary instrumentation to measure the effectiveness of different modes of counseling.

CSE Report No. 49; 41pp; Oct., 1969
CSE SIMULATED EVALUATION EXERCISE:
INSTRUCTION GUIDE

Marvin C. Alkin, Mary M. Bentzen, &
J. Eugene Grigsby III
The goal of the exercise described in this manual is to increase the flexibility of evaluators as they respond to constraints encountered in actual field conditions. Users are guided in their efforts to construct and modify evaluation designs. Specific objectives of the exercise, including scoring procedures for determining the extent to which each objective has been achieved, are also discussed.

CSE Report No. 50; 82pp; Oct., 1969
CSE SIMULATED EVALUATION EXERCISE:
MATERIALS SUPPLEMENT
Marvin C. Alkin, Mary M. Bentzen, &
J. Eugene Grigsby III
As an integral part of the Simulated Evaluation Exercise, the Materials Supplement is complementary to the Instruction Guide. It contains in simulated form informational material necessary for conducting an actual evaluation.

CSE Report No. 51; 12pp; Jan., 1969
AN EVALUATION OF HIGHER EDUCATION: PLANS
AND PERSPECTIVES ED 036 188
C. Robert Pace
Problems of evaluation methodology and analysis which must be solved in a comprehensive appraisal are raised and discussed here. The report seeks to expand the range of criteria used in judging the impact of higher education on students and to extend the range of study of environmental conditions that may have a bearing on this impact.

CSE Report No. 52; 57pp; May 1969
AN ANALYSIS OF VARIANCE FRAMEWORK FOR
MATRIX SAMPLING ED 032 648
Ken Sirotnik
This paper illustrates how a matrix sample can be conceived as an examinee-by-item analysis of variance design, the matrix sampling formulas for estimating the mean and variance of examinee scores being the formulas for estimating components of the underlying model. Several procedures are suggested to handle the problem of negative variance estimates when multiple matrix sampling is used.

CSE Report No. 53; 10pp; June, 1969
STUDENT ENTRY SKILLS AND THE EVALUATION OF
INSTRUCTIONAL PROGRAMS: A CASE STUDY
Rodney W. Skager
In experimental and traditional mathematics classes for urban minority students, teachers were observed to emphasize material in which student skills were already relatively well-developed, as compared to materials in which they were initially weak. The paper discusses the implications of this for curriculum development and evaluation methodology, and stresses the importance of mapping student entry skills before designing instructional programs.

CSE Report No. 54; 24pp; Sept., 1969
EVALUATION AND THE IMPROVEMENT OF COMPENSATORY
EDUCATIONAL PROGRAMS
Rodney W. Skager
In this paper evaluation is viewed as the collection and interpretation of systemic information about the effectiveness of alternative educational practices. Several functions of evaluation in education are described. The paper asserts that the majority of Title III programs have been primarily restricted to the program certification function and advances several proposals for improving the evaluation of compensatory education programs.

CSE Report No. 55; 52pp; Nov., 1969
MEASURING THE INSTITUTIONAL STANCE ON MATTERS
OF STUDENT CONDUCT ED 035 386
Richard Seligman
The study was undertaken to develop a method for quantifying an institution's stance on matters of student conduct. The IPCU (Institutional Procedures in Colleges and Universities) Questionnaire was completed by students and administrators selected from a group of colleges, universities, and public junior colleges. Analyses of questionnaire responses indicated that universities and junior colleges are perceived as exhibiting limited involvement in students' affairs while denominational colleges show pervasive institutional concern for students on and off campus.

CSE Report No. 56; 47pp; May, 1970
RATIONALE AND USE OF CONTENT-RELEVANT ACHIEVEMENT
TESTS FOR THE EVALUATION OF INSTRUCTIONAL PROGRAMS
Marianne Patalino
The method described in this paper for the construction, analysis, and interpretation of a test to evaluate instructional programs represents

an alternative approach to the traditional reliance on the standardized achievement test and the total scores they provide. This method leads to a content-relevant, change-sensitive instrument which was successfully applied to the evaluation of the Los Angeles Model Mathematics Project.

CSE Report No. 57; 23pp; May, 1970
THE UNINTENTIONAL MEMORY LOAD IN TESTS FOR YOUNG CHILDREN
Margaret Hubbard Jones
This paper discusses the short-term memory load of a number of standardized tests of several types in relation to the short-term memory capacity of primary school children. The characteristics of short-term memory and the rigidity of its limits are discussed and a number of factors which interfere with comprehension of language are detailed. It is argued that memory processes are of sufficient importance to be measured properly and accurately and not confounded with other skills, and that use of memory screen in tests for young children reduces both their reliability and their validity.

CSE Report No. 58; 37pp; June, 1970
MEASURING THE NORMAL AFFECTIVE STATES IN CHILDREN
Symposium presented at the 1970 Annual Convention of
the Western Psychological Association
Ralph Hoepfner, Chairman
In this symposium, four papers discuss the general topic of measurement of normal affective states in children. An attitude instrument to measure the attitudes of primary school children toward various aspects of school is described, and an interest inventory focusing on the primary school child is discussed. The symposium also discussed an attempt to simplify the task of incorporating the affective components into an existing system of goals for the school, and presents a comprehensive list of elementary school objectives compiled and condensed under 110 headings.

CSE Report No. 59; 28pp; July, 1970
DEVELOPMENT OF A SCHOOL ATTITUDE QUESTIONNAIRE
FOR YOUNG CHILDREN
Guy Strickland
This paper discusses an instrument developed to measure the attitudes of primary school children toward various aspects of school. It consists of a series of cartoon pictures. The story depicted in these pictures is described orally. Responses are made directly on the test page. This format is used to avoid the usual problems of most attitude measures that are administered to primary children. The factorial structure, reliability, and validity of the scales are discussed, along with implications for the instrument's implementation.

CSE WORKING PAPERS	ERIC Numbers
1. Gordon, C. W. Preliminary Evaluation Report on the Los Angeles City Schools SB 28 Demonstration Program in Mathematics.	ED 035 784
2. Jones, M. Reliability of Coding of the System for the Analysis of Classroom Communication.	ED 036 856
3. Wittrock, M.C. & Jones, M.H. Transfer of Rules: The Need for Tests of Transfer in the Evaluation of Instruction.	ED 035 946
4. Alkin, M.C. & Woolley, D.C. A Framework for Evaluation of TESOL Programs.	
5. Olivier, K., Jr. Data Collection and Reduction Procedures Used for the System for Analysis of Classroom Communication (SACC).	ED 035 945
6. Wittrock, M.C. & Jones, M.H. Transfer of Abstract and Specific Instruction in Children's Learning.	
7. Kennedy, G. The Language of Tests for Young Children.	ED 035 781
8. Lindman, E. L. Net-Shift Analysis for Comparing Raw Score Distributions.	
9. Alkin, M. C. Objectives and Objective-Based Measures in Evaluation.	
10. Alkin, M. C. A Review of the Evaluation of the Follow Through Program.	

REPRINT SERIES

1. Pace, C. Robert. "The Measurement of College Environments," from Organizational Climate: Explorations of a Concept. Ranato Tagiuri and George H. Litwin (Eds.), Boston: Graduate School of Business Administration, Harvard University, 1968, 129-147.

2. Feshbach, Norma D., & Beigel, Astrid. "The Use of the Semantic Differential in Measuring Teacher Personality and Values," from Educational and Psychological Measurement, 1968, 28, 923-929.

3. Feshbach, Norma D. "Student Teacher Preferences for Elementary School Pupils Varying in Personality Characteristics," from Journal of Educational Psychology, 1969, 60 (2), 126-132.

4. Robinson, Lora., & Seligman, Richard. "A Scale for Measuring Campus Morale," from Journal of Educational Measurement, 1969, 6 (2), 109-110.

5. Hoffenberg, Marvin. "Program Budgeting in Education: Some Organizational Implications," from Strategies of Educational Planning, Richard H.P. Kraft, (Ed.), Educational Systems Development Center, College of Education, The Florida State University, Tallahassee, Florida, 1968, 184-213.

6. Feshbach, Norma D., & Devore, Geraldine. "Teaching Styles in Four-Year-Olds," from Child Development, 1969, 40 (1), 183-190.

7. Popham, W. James. "Implications of Criterion-Referenced Measurement," from Journal of Educational Measurement, 1969, 6 (1), 1-9.

8. Popham, W. James. "Focus on Outcomes: A Guiding Theme of ES '70 Schools," from Phi Delta Kappan, 1969, 51 (4), 208-210.

9. Boiarsky, Carolyn. "Help for Your District: How to Set, and Meet, Learning Goals," from The American School Board Journal, 1969, 157 (6), 12-13

10. Baker, Eva L. "Experimental Assessment of the Effects of the Probe System," a symposium presentation at the American Educational Research Association Annual Meeting, Minneapolis, Minnesota, March, 1970.

11. Baker, Eva L. "The Effects of Manipulated Item MRiting Constraints on the Homogeneity of Test Items," a paper presented at the Annual Conference of the National Council for Measurement in Education, Minneapolis, Minnesota, March, 1970.

12. Hoffenberg, Marvin, & Alkin, Marvin C. "Application of Leontief Input-Output Analysis to School Districts Budgeting," a symposium presentation at the American Educational Research Association Annual Meeting, Minneapolis, Minnesota, March, 1970.

13. Seligman, Richard. "Student and Administrator Perception of Campus Discipline," a symposium presentation at the American Educational Research Association Annual Meeting, Minneapolis, Minnesota, March, 1970.

14. Popham, W. James. "Indices of Adequacy for Criterion-Referenced Items," a symposium presentation at a Joint Session of the National Council for Educational Measurement and the American Educational Research Association, Minneapolis, Minnesota, March, 1970.

15. Popham, W. James. "The Instructional Objectives Exchange: Progress and Prospects," a symposium presentation at the American Educational Research Association Annual Meeting, Minneapolis, Minnesota, March, 1970.

16. Skager, Rodney. "Objective Based Evaluation: Macro-Evaluation," a symposium presentation at the American Educational Research Association Annual Meeting, Minneapolis, Minnesota, March, 1970.

17. Hoepfner, Ralph. "Intellectual Aptitude Involvement in Thinking Skills," from the Journal of Research and Development in Education, 1969, 3 (1), 43-51.

18. Hoepfner, Ralph. "Evaluating a High School 'Gifted' Program," a paper delivered at the National Council of Measurement in Education Annual Convention in Minneapolis, Minnesota, March, 1970.

19. Popham, W. James. "Instructional Objectives 1960-1970," an address presented at the Eighth Annual Convention of the National Society for Programmed Instruction, Anaheim, California, May, 1970.

20. Alkin, Marvin C. "Educational Needs for the 1970's and Beyond," from Needs of Elementary and Secondary Education for the Seventies, Committee on Education and Labor, House of Representatives. U.S. Government Printing Office, Washington, D.C., March, 1970.

21. Pace, C. Robert. "An Evaluation of Higher Education: Plans and Perspectives," from the Journal of Higher Education, 1969, 40 (9), 673-681.

EVALUATION COMMENT ARTICLES

Vol. 1, No. 1 - "A New Role in Education: The Evaluator," by Garth Sorenson; "Cognitive Skills: A Consideration in Evaluating Instructional Effects," by Rodney W. Skager.

Vol. 1, No. 2 - "Learning for Mastery," by Benjamin S. Bloom.

Vol. 1, No. 3 - "Matrix Sampling," by T.R. Husek and Ken Sirotnik.

Vol. 1, No. 4 - "The Evaluation of Instruction: Cause and Effect Relations in Naturalistic Data," by M.C. Wittrock.

Vol. 2, No. 1 - "Evaluation Theory Development," by Marvin C. Alkin; "Different Kinds of Evaluation and their Implications for Test Development," by T.R. Husek.

Vol. 2, No. 2 - "Evaluating Tests in Terms of the Information They Provide," by Stephen Klein; "Objective Based Evaluation: Macro-Evaluation," by Rodney W. Skager.

ADDITIONAL CENTER PUBLICATIONS

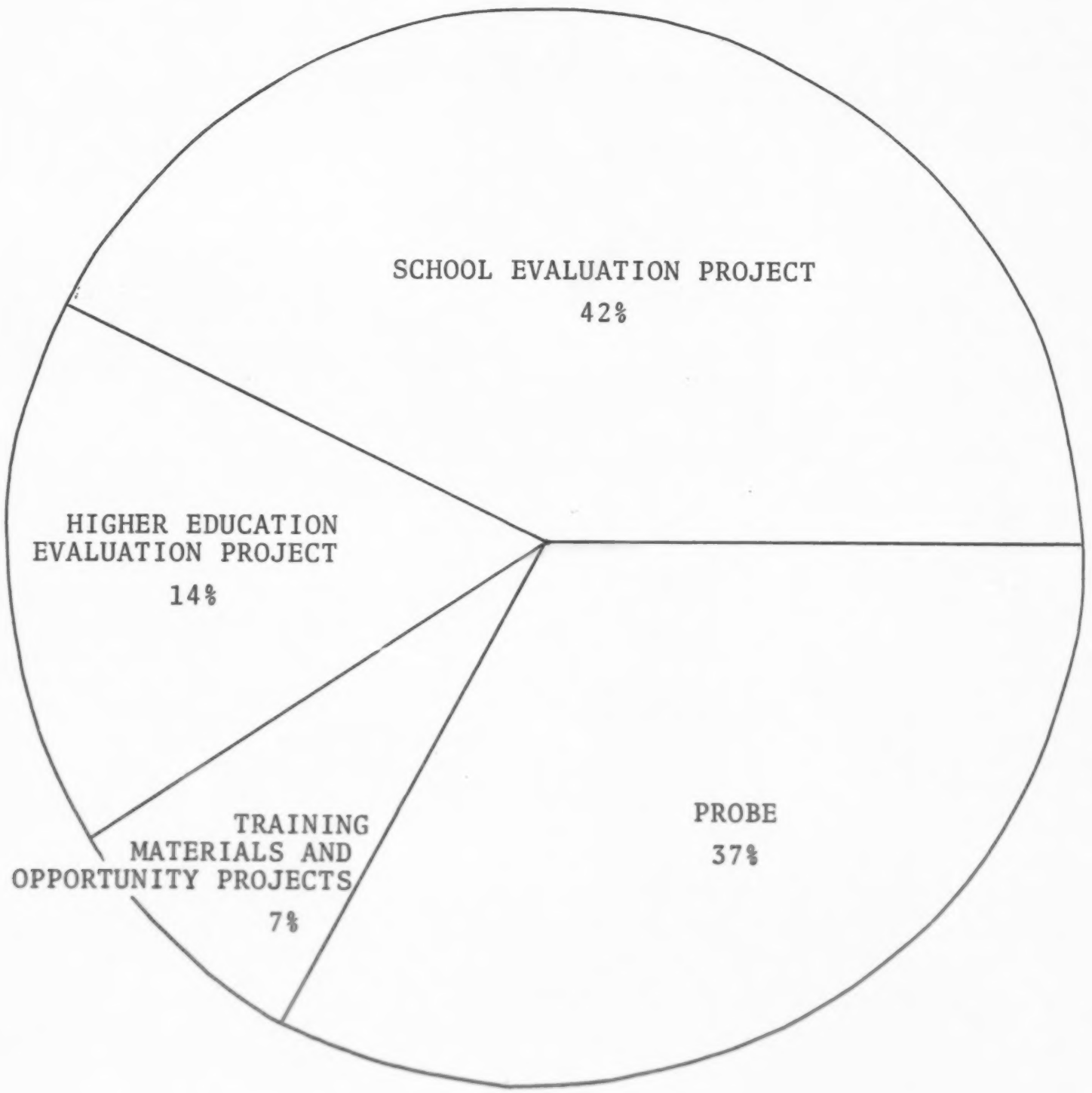
Alkin, M.C., Baker, E.L., Hunter, M., McIntire, R.G., Popham, W.J., and Skager, R.W., (guiding committee), Instructional Objectives Exchange: Objective Collections, sixteen collections, Los Angeles: Center for the Study of Evaluation, 1969.

Hoepfner, R., Strickland, G., Stangel, G., Jansen, P., and Patalino, M. CSE Elementary School Test Evaluations. Los Angeles: Center for the Study of Evaluation, UCLA, 1970.

APPENDIX D

CSE Operating Funds
Fiscal Year 1970
Percentage Allocation by Project

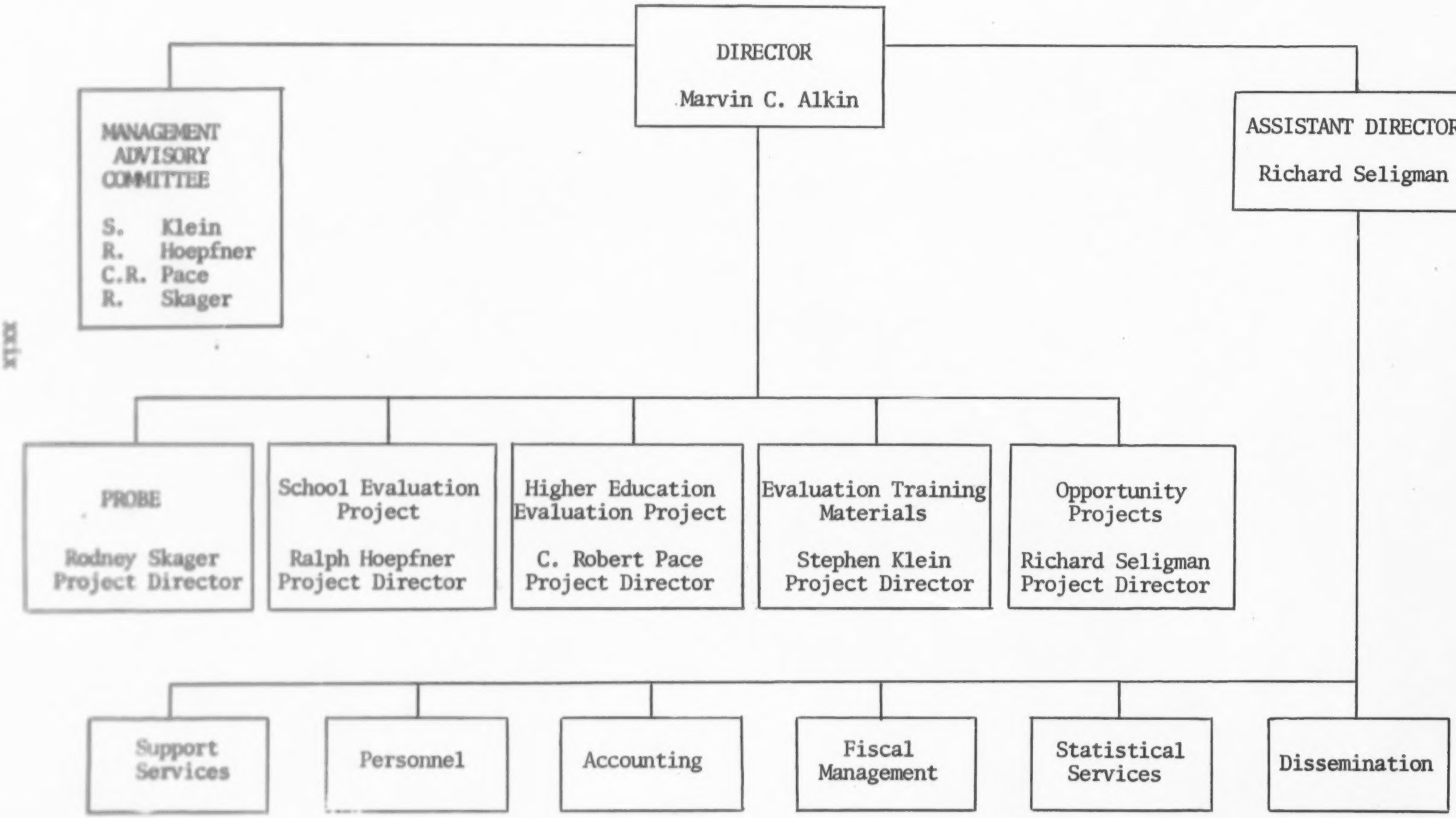
CSE OPERATING FUNDS
FISCAL YEAR 1970
PERCENTAGE ALLOCATION BY PROJECT



APPENDIX E

CSE Administrative Structure

CSE ADMINISTRATIVE STRUCTURE



APPENDIX F

Regional Laboratories and R & D Centers Currently in Operation

REGIONAL LABORATORIES AND R & D CENTERS CURRENTLY IN OPERATION [1]

1) Reproduced by permission of Phi Delta Kappan (See Francis S. Chase, "R&D in the Remodeling of Education", Phi Delta Kappan, 51(6), 1970, p. 301.)